Surveillance of SARS-CoV-2 at the Huanan Seafood Market

Running title:

Prevalence of SARS-CoV-2 in Huanan Seafood Market

William J. Liu,[1]# Peipei Liu,[1]# Wenwen Lei,[1]# Zhiyuan Jia,[1]# Xiaozhou He,[1]# Weifeng Shi,[2]# Yun Tan,[3]# Shumei Zou,[1] Gary Wong,[4] Ji Wang,[1] Feng Wang,[1] Gang Wang,[1] Kun Qin,[1] Rongbao Gao,[1] Jie Zhang,[1] Min Li,[1] Wenling Xiao,[1,5] Yuanyuan Guo,[1] Ziqian Xu,[1] Yingze Zhao,[1] Jingdong Song,[1] Jing Zhang,[1] Wei Zhen,[1] Wenting Zhou,[1] Beiwei Ye,[1] Juan Song,[1] Mengjie Yang,[1] Weimin Zhou,[1] Yuting Dai,[3] Gang Lu,[3] Yuhai Bi,[6] Wenjie Tan[1], Jun Han[1], George F. Gao,[1,6] Guizhen Wu[1]

[1] NHC Key Laboratory of Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention (China CDC), Beijing 102206, China;

[2] Key Laboratory of Emerging Infectious Diseases in Universities of Shandong, Shandong First Medical University, and Shandong Academy of Medical Sciences, Tai'an 271000, China;

[3] Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine, Ruijin Hospital Affiliated to Shanghai Jiao Tong University (SJTU) School of Medicine, Shanghai 200020, China;

[4] CAS Key Laboratory of Molecular Virology & Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences (CAS), Shanghai 200031, China;

[5] School of Laboratory Medicine and Life Sciences, Wenzhou Medical University, Wenzhou 325035, China;

[6] CAS Key Laboratory of Pathogen Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing 100101, China.

# Contributed equally.

Correspondence:
Guizhen Wu: wugz@ivdc.chinacdc.cn, ORCID 0000-0003-2778-4290
George F. Gao: gaofu@chinacdc.cn, ORCID 0000-0002-3869-615X
William J. Liu: liujun@ivdc.chinacdc.cn, ORCID 0000-0003-3605-4070

Abstract

Emerging in December 2019, coronavirus disease 2019 (COVID-19) eventually became a pandemic and posed a tremendous threat to global public health. However, the origins of SARS-CoV-2, the causative agent of COVID-19, remain to be determined. It has been reported that a certain number of the early case clusters had a contact history with the Huanan Seafood Market. Therefore, surveillance of SARS-CoV-2 within the market is of vital importance. Herein, we presented the SARS-CoV-2 detection results of 1380 samples collected from the environment since 1st Jan and animals since 18th Jan within the market in early 2020. By SARS-CoV-2-specific RT-qPCR, 73 environmental samples tested positive for SARS-CoV-2 and three live viruses were successfully isolated. The viruses from the market shared nucleotide identity of 99.99% to 100% with the human isolate HCoV-19/Wuhan/IVDC-HB-01/2019. The A lineage (8782T and 28144C), as the likely ancestral SARS-CoV-2 lineage, was found in an environmental sample. No virus was detected in the animal swabs covering 18 species of animals in the market. The RNA-seq analysis of SARS-CoV-2 positive/negative environmental samples showed the abundance of different vertebrata genera. In summary, this study provided convincing evidence of the prevalence of SARS-CoV-2 in the Huanan Seafood Market during the early stage of COVID-19 outbreak.

Keywords:

COVID-19, SARS-CoV-2, Huanan Seafood Market, origin, high-throughput sequencing, virus isolation, sewage

Infections with novel human coronavirus 2019 (HCoV-19) [1,2], named as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses (ICTV) [3], can result in coronavirus disease 2019 (COVID-19), named by World Health Organization (WHO) characterized by various clinical outcomes from asymptomatic infections to severe pneumonia and even death [4,5]. Globally, as of Feb 28th 2023, over 758 million confirmed cases and over 6.8 million deaths have been reported (covid19.who.int).

Early human cases with COVID-19 were first reported in late December 2019 from Wuhan, China, with pneumonia of unknown etiology (PUE), and a majority of them were found to be linked to the Huanan Seafood Market (HSM) in Wuhan [4,6], where various animal meats, exotic seafoods and live animals were available for purchase. Therefore, the HSM has been suspected to be the source of the COVID-19 pandemic [7]. However, alternative hypothesis that some individuals outside the market had brought the virus into the market through human-to-human transmission or cold chain could not be ruled out, considering some of the early cases without epidemiological link to the market [6,8]. Yet, more scientific evidence and further studies are indeed needed.

Considering the findings that SARS-CoV-2 has high similarities with a few coronaviruses derived from bats in Asian countries including China, Laos, Japan, Cambodia and Thailand, scientists have proposed that bats might be the original source of SARS-CoV-2 [1,9-14]. However, it is still a mystery whether another animal might act as an intermediate host to have facilitated the virus spillover from bats to humans [15,16]. One of such important findings was the discovery of SARS-CoV-2 related coronaviruses from pangolins that harbored highly similar receptor binding domain (RBD) with that of SARS-CoV-2 in the viral spike (S) protein [17-19]. Despite that pangolins might be involved in the ecology of coronaviruses, whether the pangolins are the intermediate hosts is yet unknown, given the current data[20]. A recent study documented the animal species in the HSM between May 2017 and November 2019

and noted that no pangolins or bats, but some hypothesized Sarbecovirus-susceptible animals such as Raccoon dogs were present [21]. Thus far, the origins of SARS-CoV-2 [22,23] and the role of the HSM in the origins and spread of SARS-CoV-2 remain unclear. The data from the HSM may provide important information in this field.

The HSM is located in the Jianghan District, a downtown of Wuhan, the capital city of Hubei Province, and is approximately 800 m far from Hankou Railway Station, a major railway travel hub. It occupies >50,000 m^2, with 678 stalls located close to each other in an extremely crowded condition (Fig. 1A). The market is separated into two zones, East and West Zones, with seafood and animals mainly sold in West Zone and livestock meat in East Zone. Among the 678 stalls of the market, 10 domesticated wildlife animal-selling stalls (1.5%) were identified, located in the south-western corner of West Zone (8/10) and the north-western corner of East Zone (2/10), respectively (Fig. 1A). According to sale records, during late December 2019, animals or animal products were sold in these 10 animal stalls. Animals included snakes, avian species (chickens, ducks, gooses, pheasants and doves), Sika deer, badgers, rabbits, bamboo rats, porcupines, hedgehogs, salamanders, giant salamanders, bay crocodiles and Siamese crocodiles, etc., among which snakes, salamanders and crocodiles were traded as live animals (described in detail in the Report of WHO-convened global study of origins of SARS-CoV-2 [24]).

The market was closed in the morning of January 1st, 2020, shortly after the identification of the PUE. At the same time, in order to investigate the potential introduction of SARS-CoV-2 into the market, Chinese Center for Disease Control and Prevention (China CDC) dispatched an epidemiological team, together with experts from Hubei Provincial CDC and Wuhan Municipal CDC, to the HSM to collect environmental samples in the early morning of January 1st, 2020 (Fig. 1B). From January 1st, 2020, to March 2nd, 2020, a total of 923 environmental samples from different locations within and around the market and 457 animal samples including

animal bodies, stray animals and their feces, were collected, with some stray animals sampled until March 30^{th} (Extended Data Tables 1, 2, 3 and Supplementary Table 1). This may reflect the profile of SARS-CoV-2 contamination in the market during the early phase of the outbreak. After the closure of the market, the outside surface of the rolling shutter doors of the stalls and the corridors were disinfected (with 1% bleach mixed with water) throughout January and February 2020. The goods inside the stalls were completely cleared and disinfected until early March 2020.

Out of the 923 environmental samples collected in and around the HSM, 73 were found to be positive for SARS-CoV-2 with positive rate of 7.9% through the nucleic acid test (NAT), with cycle threshold (CT) values of real-time polymerase chain reaction (PCR) ranging from 23.9 to 41.7 (Table 1). Among the 828 samples inside the HSM, 64 samples (7.7%) were positive. For the 14 samples from warehouses related to the HSM, five tested positive. Among the 51 samples from sewerage wells (Supplementary Table 1) in the surrounding areas outside the HSM, three were tested positive (Table 1). Notably, one sample (Env_0601) out of the 30 environmental samples, a floor surface swab collected from Dongxihu Market in Wuhan on January 22nd, 2020, was also tested positive (Table 1, Extended Data Table 4).

For the 64 SARS-CoV-2 positive samples inside the HSM, 87.5% (56/64) were collected in the West Zone of the market, in particular streets from no. 1 to 8, with 71.4% (40/56) positive samples identified herein (Fig. 1A). Of the 110 samples collected from sewers or sewerage wells in the market, 24 samples were positive for SARS-CoV-2 nucleic acid. All the four sewerage wells in the market tested positive. During the onsite investigation of the overground drainage pathway in the HSM, we found that the wastewater in the overground drainage was led into the underground drainage inside the market and then flow into the wells on the edge of the market. We then did a spot-check sampling across all the overground drainages according to the principles described in the Methods (Extended Data Fig. 1). In fact, the excreta of upper

respiratory tract of the patients and the potential animal waste would be mixed together into the overground drainage. Thus, these data suggested that either the infected people and/or animals in the market contaminated the sewage or that the contaminated sewage might have further played a role in the virus transmission within the case cluster in the market.

The merchants' activities were assessed against the NAT results of the environmental samples. The sampling covered 19.8% (134/678) of the vendors in the market (95% confidence interval (CI): 16.8-23.0%). Of the positive samples, 44 were distributed among 21 vendors in the market, 19 of whom were located in the West Zone and the remaining two located in the east area (Fig. 1A). Some vendors sold more than one product type. While the results provided some indication of association of cases with different products, no significant differences were observed between different vendors, including poultry (22%, 8/37: 95% CI: 9.8-38.2%), cold-chain products (18.4%, 16/87, 95% CI: 10.9-28.1%), aquatic products (17.8%, 13/73, 95% CI: 9.8-28.5%), livestock (14%, 5/36: 95% CI: 4.7-29.5%), seafood products (11%, 6/56: 95% CI: 4-21.9%), wildlife products (11%, 1/9: 95% CI: 0.3-48.2%), and vegetables (25%, 2/8: 95% CI: 3.2-65%) (Extended Data Fig. 2, Extended Data Table 5). The positive vendors with multiple product types suggested that SARS-CoV-2 might have been circulating in the market, especially the West Zone, for a while in December 2019, leading to an extensive distribution of the virus within the market, which might have been facilitated by the crowded buyers and the contaminated environment.

The 457 animal samples included 188 individuals belonging to 18 species (with some stray animals sampled until March 30th) (Table 2). The sources of the samples included unsold goods kept in refrigerators and freezers in the stalls of the HSM, and goods kept in warehouses and refrigerators related to the HSM. Three Chinese giant salamanders, which were found in a fish tank, were alive and swab samples were collected and tested. Samples from stray animals in the market were also collected, i.e. swab samples from

10 stray cats, 27 cat feces, one dog, one weasel, and 10 rats. All the 457 animal samples were tested negative for SARS-CoV-2 nucleic acid.

To determine whether there was live virus in the HSM, we inoculated 27 SARS-CoV-2 positive environmental samples collected on January 1st, 2020, into cell lines, including Vero E6 and Huh7.5 cells. Cytopathic effects (CPE) were observed 3 days post inoculation with sample Env_0313 on Vero E6 cells. CPE was also observed 5 days post inoculation on Huh7.5 cells. The electron micrographs of Vero E6 cells after 5 days of post inoculation showed that virus particles were present in both the supernatant and the cells. Negative-stained virus particles and ultra-thin cultured cell sections showed typical coronavirus morphology (Fig. 2). Totally, live viruses were isolated from samples Env_0313, Env_0354 and Env_0126, which were the only three samples with CT values <30 in the NAT. Env_0354 and Env_0126 were swab samples of the ground and Env_0313 were swab samples of the wall. Notably, samples Env_0313 and Env_0354 were from the stalls with confirmed patients. All the results of successful virus isolation and the CT values of the original samples revealed the existence of live SARS-CoV-2 with high titers in the environment of the HSM. We did not perform virus isolation based on the samples collected from later time points due to the high CT values.

During later sampling in the HSM on February, we specially collected some samples to investigate the virus RNA persistence in the Market. Some of the samples were still tested positive, especially in the sewage well and even on the walls (Table 1). For the sample Env_0838 on the wall collected on February 20th, 2020, the results of the three PCR targets were 32.59/“-”/37.34, accordingly. This result is reasonable considering the degradation of the viral genome and some of the PCR target cannot be detected. However, the results also indicate a long persistence of the virus RNA in the environment.

We further performed high-throughput sequencing (Extended Data Table 7) and successfully obtained seven complete or near complete SARS-CoV-2 genome sequences, including three sequences from three environmental samples (Env_0313, Env_0354 and Env_0020), and four sequences from cell supernatants of Env_0313, Env_0354 and Env_0126 (Fig. 3, Extended Data Table 7). The genome sequences of three environmental samples, Env_0126, Env_0313 and Env_0354, were found to be completely identical to the reference strain HCoV-19/Wuhan/IVDC-HB-01/2019 (IVDC-HB-01, GISAID accession number: EPI_ISL_402119) and the human strain Wuhan-Hu-1 (GenBank: NC_045512) (Fig. 3A). The genome sequence of the isolated virus from environmental sample Env_0354 had two synonymous mutations compared to HCoV-19/Wuhan/IVDC-HB-01/2019, with sequence identity of 99.99% (Fig. 3A). Therefore, the SARS-CoV-2 sequences from environmental samples were highly similar to the clinical strains obtained during the early stage of the COVID-19 outbreak.

Previously, SARS-CoV-2 has been proposed to be classified into two major lineages based on the two highly-linked single nucleotide polymorphisms (SNPs): A lineage (8782T and 28144C, or S lineage in another nomenclature of SARS-CoV-2) and B lineage (8782C and 28144T, or L lineage). It has been proposed that A/S lineage most likely is the ancestral lineage, because all of the SARS-CoV-2 related coronaviruses from bats and pangolins possessed 8782T and 28144C [25,26], while Pekar et al. also presented a possibility that both lineages represent separate introduction events[27]. Phylogenetic analysis revealed that most of the environmental strains belong to the B/L lineage and they cluster together with the human strains circulating in the early stage of the pandemic (Fig. 3B, Supplementary Fig. 1). However, the environmental sample Env_0020 falls within the A/S lineage in the tree (Fig. 3B), with the confirmation by the high number of reads mapped to positions 8782 and 28144 in Env_0020 (Extended Data Table 8). However, it should be noted that the genome of Env_0020 is of low quality and there are many discontinuous gaps in the assembled genome. Indeed, though it is difficult to root the SARS-CoV-2 phylogenetic tree, our analysis indicated

that the environmental viruses cluster together with the human strains circulating in the early stage of the pandemic.

We conducted RNA-seq analysis using 60 SARS-CoV-2 PCR-positive and 112 SARS-CoV-2 PCR-negative environmental samples from the HSM (Fig. 4A and Extended Data Table 6). Bacteria were the most abundant species in almost all samples and mammal species could be found in most samples, which fit the feature of samples collected from the environment (Fig. 4B and Supplementary Table 2). We further classified the vertebrate genera using the barcode of life data system. The *Gallus*, *Homo*, *Anas*, *Sus*, *Bos*, and *Canis* could be detected in most samples (Fig. 4C and Supplementary Table 3), which was in accordance with the environmental feature of the seafood markets in China. In addition, The most different genus in the bacteria domain belonged to *Flavobacteriia* class (higher in SARS-CoV-2 PCR-positive samples) and *Clostridia* class (higher in SARS-CoV-2 PCR-negative samples) (Supplementary Table 4). We analyzed the mammal genera in all sequenced samples with the kranken2 (detailed in the method) using different thresholds (Fig. 4D and Supplementary Fig. 3). A total of 70 mammal genera, which existed in more than 2% samples, were identified with a threshold of 100 reads per million (Fig. 4D and Supplementary Table 5).

Particularly, we analyzed three samples (the B5, F13, and F54) collected on 1st Jan 2020 with high levels of SARS-CoV-2 (Ct value <30) (Fig. 4E). The identified mammal genera in the F13 and F54 were related to species in the general food market, such as *Homo*, *Ovis*, *Bos*, *Canis*, *Sus*, and *Felis*. Many mammal genera were observed in the B5 sample, but the most abundant mammal genera were also related to the general food market, including *Bos* (77.30%), *Ovis* (19.91%), *Homo* (0.77%), and *Bubalus* (0.57%). Although *Pipistrellus*, a kind of bat, can also be found in this sample, but the relative ratio was low (0.002%), and the *Pipistrellus* is unlikely the kind of bat that host SARS-CoV-2 as previously reported. Meanwhile, the *Lutra*, another suspected host for SARS-

CoV-2, was also observed in this sample, but the relative ratio was extremely low (0.001%). It should be noted that mammals share some sequences in genomes. Thus, the more mammal reads detected, the more species will be found, and those identified species/genera with relatively low supporting reads were likely unreliable in those samples with abundant detected mammal reads. Moreover, we also noted that only *Homo*, *Ovis*, *Bos*, and *Sus* reads but not species related to wildlife were found in the Env_0020 samples, the one that falls within the A/S lineage in the tree.

In addition, we illustrated the top-ranked genera in four areas of the market, where multiple SARS-CoV-2 PCR-positive samples were detected. As shown in Fig. 4F, the top-ranked genera in these areas were homo or other genera generally existed in food markets. We also noted the *Nyctereutes* could be found in the vender 25 of the street 8, while *Atelerix* and *Erinaceus* could also be found in vender 15-17 of street 7 (Fig. 4F). However, these genera existed in both SARS-CoV-2 PCR positive and negative samples (Supplementary Table 2, 3), not to mention the positive ratio was much higher in SARS-CoV-2 PCR negative samples.

Also, we checked samples that might relate to wildlife, such as samples collected in the defeathering machine and the visible blood spot. Top abundant mammal genera of the defeathering machine sample (Env_0584) was *Canis* (Extended data Fig 3). Top abundant mammal species of the visible blood spot (Env_0262) were *Bos*, *Sus*, *Ovis* and *Bison*, accordingly (Extended data Fig 3). Additionally, we plotted the distribution of some genera of concern, including *Myotis*, *Erinaceus*, *Mustela*, *Nyctereutes*, *Rhizomys*, *Meles*, and *Melogale*. Most of these samples were distributed in the western district of the market (Extended data Fig 4), where they sold wildlife products. Although the enriched areas of SARS-CoV-2 PCR-positive samples were nearby, the distribution locations of *Homo*, *Sus*, *Bos*, *Gallus* and *Anas* were also dominated in this area. The repeated sampling of the locations with PCR-positive results may contribute some bias on the distribution analyses of enriched areas of SARS-CoV-2 PCR-positive

samples. Moreover, we plotted the distribution of mammal genera in those SARS-CoV-2 positive samples with high abundance of genera related to wildlife, such as the Env_0576 (*Nyctereutes* enriched), Env_0807 (*Lariscus* enriched), Env_0809 (*Erinaceus* enriched), and Env_0585 (*Erinaceus* enriched) (Extended Data Fig. 3). We noted multiple wildlife-related genera could be identified, and the relative ratio was low when using the kraken2 methods, except for the *Atelerix* and *Erinaceus*, genera that have been reported in the WHO report.

Of particular note was the difference in the results from PCR and NGS. In fact, among the 60 SARS-CoV-2 PCR-positive samples for RNA-seq analysis, 39 samples tested negative by NGS (no SARS-CoV-2 reads at all) (65.0%), including the samples Env_0262. For these NGS-negative samples, the CT values ranged from 31.80 to 37.44. Since the RT-PCR detection assay employed in the very early stage of the pandemic was not formally verified, we believe that there might be some false positives in the PCR detection results in this study. Meanwhile, we also found the SARS-CoV-2 reads could also be detected in a portion of SARS-CoV-2 PCR negative samples (15.2%), which might be caused by the degradation of SARS-CoV-2 within the PCR target region.

In summary, SARS-CoV-2 RNA was detected in stalls in the West Zone of the HSM, suggesting the prevalence of SARS-CoV-2 in the market. Thus, although the origin of the virus cannot be determined from all the analyses available so far, the market might at least have acted as an amplifier due to the high number of visitors every day, causing many initially identified infection clusters in the early stage of the outbreak [24]. In addition, live SARS-CoV-2 viruses also existed in the environmental samples. More importantly, this is the first time for SARS-CoV-2 virus isolation from environmental samples. No SARS-CoV-2 was detected in the animal samples from the HSM. However, it should also be noted that samples analyzed in the current study were somehow biased, and wildlife-related vendors and early case-related vendors were prioritized for sample collection. Actually, the gene barcode analysis of animal species in the study showed

Myotis, *Nyctereutes* and *Melogale* that have been recognized as potential host species of Sarbecoviruses were mostly detected within the SARS-CoV-2 PCR negative environment samples.

Meanwhile, recent reports traced the outbreak back to the HSM and suspected that the market sold live animals as recently as 2019, by compiling information reported by various sources, including the WHO-China Joint Report and social media, etc[28]. Another report hypothesized that SARS-CoV-2 was spilled over from animals to humans at least twice in November or December 2019, even the raccoon dog was suspected as the intermediate animal[27]. Nevertheless, the new evidence pinpointing to the market does not support such a hypothesis and an alternative hypothesis could be further studied [29]. Our study confirmed the existence of raccoon dog and other hypothesized/potential SARS-CoV-2 susceptible animals through gene barcode within the market before the closure. However, these environmental samples cannot prove the infection of the animals. Furthermore, even if the animals were infected, it could not still rule out that the human-to-animal transmission occurred, considering the sampling time was at least one month after the human-to-human transmission within the market. Thus, the possibility of potential introduction of the virus through human or cold chain product into the market cannot yet be ruled out.

Definitely, more work involving international coordination is needed to investigate the potential origins of SARS-CoV-2 [24]. Surveillance of wild animals using a viromic approach should be enhanced to explore the potential natural and intermediate hosts for SARS-CoV-2 [7,30], if any, which would help to prevent future pandemics caused by animal-origin coronaviruses or alike, with a spillover event.

Acknowledgements

We gratefully acknowledge experts from Wuhan City, Hubei Province and across China who contributed to the study. We gratefully acknowledge the following experts for their

invaluable contributions during this study: Prof. Naiying Mao and Prof. Yu Lan from National Institute for Viral Disease Control and Prevention, China CDC; Prof. Huaiqi Jing, and Prof. Qiyong Liu from National Institute for Communicable Disease Control and Prevention, China CDC; Prof. Lei Xu from Tsinghua University; Prof. Yongzhong Jiang, Junqiang Xu, Prof. Xixiang Huo and Dr. Bo Yu from Hubei Provincial CDC; Prof. Yan Xiong from Wuhan Municipal CDC; Prof. Juan Li from Shandong First Medical University; Prof. Weijun Chen and Dr. Honglong Wu from BGI PathoGenesis Pharmaceutical Technology. In addition, we also thank the work and suggestions of the joint team scientists of WHO-convened Global Study of Origins of SARS-CoV-2: China Part. Prof. William J. Liu is supported by the Excellent Young Scientist Program of the National Natural Science Foundation of China (NSFC, 81822040). Dr. Yun Tan's bioinformatics analyses in the study were supported by the ASTRA computing platform in the National Research Center for Translational Medicine (Shanghai) and the Pi computing platform in the Center for High Performance Computing at Shanghai Jiao Tong University.

Author contributions

The study was designed by G-Z.W., W.J.L and G.F.G. The onsite epidemiological survey and sample collection by W.J.L., W.L, Z.J., X.H., J.W., F.W., G.W., K.Q., R.G., J.Z., M.L. W.X. and G.F.G. The nucleic acid extraction and RT-PCR were performed by W.J.L., P.L., W.L, Z.J., X.H., J.W., F.W., K.C. and G.W. Next generation sequencing was performed by W.J.L., P.L., W.L, Z.J., X.H., J.W., F.W., G.W., and W.Z. Complete genome sequencing and analyses were performed by P.L., W.Z., W.S. and W.J.L. The virus isolation was performed by P.L., S.Z., W.Z., W.L., J.S. and Z.X. Data analyses were performed by W.J.L., P.L., Z.J., X.H., W.S., Y.T., S.Z., J.W., F.W., G.W., Y.G., Z.X., Y.Z., J.S., Jing Z., W.Z., W-T.Z., B.Y., J.S., M.Y., W-M.Z., Y.D., G.L., Y.B., W.T., and J.H. The manuscript was written by W.J.L., P.L., W.S., Y.T., G.W., G.F.G. and G-Z.W.

Competing interest declaration

No competing interest exists.

Data availability

All the raw sequencing data have been uploaded onto the GISAID (China CDC Weekly, 2021, DOI: 10.46234/ccdcw2021.255). The list of accession codes in Extended Data Table 6 and 7. The raw sequence data reported in this paper have also been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics, 2021, DOI: 10.1016/j.gpb.2021.08.001) in National Genomics Data Center (Nucleic Acids Res, 2022, DOI: 10.1093/nar/gkab951), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences that are publicly accessible at https://ngdc.cncb.ac.cn/gsa (GSA: CRA010170). The viral genomes reported in this paper have been deposited in the GenBase in National Genomics Data Center (Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation under accession number C_AA002295.1 to C_AA002301.1 that is publicly accessible at nttps://ngdc.cnc.ac.cn/genbase. Raw sequence data was deposited into NCBI BioProject under accession number PRJNA948658 (http://www.ncbi.nlm.nih.gov/bioproject/948658) and in China National Microbiology Data Center (NMDC) with accession numbers NMDC10018366 (https://nmdc.cn/resource/genomics/sample/detail/NMDC10018366).

References

1 Tan, W. *et al.* A novel coronavirus genome identified in a cluster of pneumonia cases - Wuhan, China 2019-2020. *China CDC Wkly* **2**, 61-62, doi: 10.46234/ccdcw2020.017 (2020).

2 Jiang, S. *et al.* A distinct name is needed for the new coronavirus. *Lancet* **395**, 949, doi:10.1016/S0140-6736(20)30419-0 (2020).

3 Coronaviridae Study Group of the International Committee on Taxonomy of, V. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **5**, 536-544, doi:10.1038/s41564-020-0695-z (2020).

4 Wang, C., Horby, P. W., Hayden, F. G. & Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **395**, 470-473, doi:10.1016/S0140-6736(20)30185-9 (2020).

5 Zhu, N. *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* **382**, 727-733, doi:10.1056/NEJMoa2001017 (2020).

6 Li, Q. *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus-infected Pneumonia. *N Engl J Med* **382**, 1199-1207, doi:10.1056/NEJMoa2001316 (2020).

7 Daszak, P., Olival, K. J. & Li, H. A strategy to prevent future epidemics similar to the 2019-nCoV outbreak. *Biosaf Health* **2**, 6-8, doi:10.1016/j.bsheal.2020.01.003 (2020).

8 Chen, N. *et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507-513, doi:10.1016/S0140-6736(20)30211-7 (2020).

9 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273, doi:10.1038/s41586-020-2012-7 (2020).

10 Murakami, S. *et al.* Detection and characterization of bat Sarbecovirus phylogenetically related to SARS-CoV-2, Japan. *Emerg Infect Dis* **26**, 3025-3029, doi:10.3201/eid2612.203386 (2020).

11 Wacharapluesadee, S. *et al.* Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat Commun* **12**, 972, doi:10.1038/s41467-021-21240-1 (2021).

12 Zhou, H. *et al.* A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr Biol* **30**, 2196-2203 e2193, doi:10.1016/j.cub.2020.05.023 (2020).

13 Zhou, H. *et al.* Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell* **184**, 4380-4391 e4314, doi:10.1016/j.cell.2021.06.008 (2021).

14 Li, J., Lai, S., Gao, G. F. & Shi, W. The emergence, genomic diversity and global spread of SARS-CoV-2. *Nature* **600**, 408-418, doi:10.1038/s41586-021-04188-6 (2021).

15 Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574, doi:10.1016/S0140-6736(20)30251-8 (2020).

16 Wang, J. *et al.* Individual bat viromes reveal the co-infection, spillover and emergence risk of potential zoonotic viruses. *bioRxiv*, doi:10.1101/2022.11.23.517609 (2022).

17 Lam, T. T. *et al.* Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **583**, 282-285, doi:10.1038/s41586-020-2169-0 (2020).

18 Xiao, K. *et al.* Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* **583**, 286-289, doi:10.1038/s41586-020-2313-x (2020).
19 Niu, S. *et al.* Molecular basis of cross-species ACE2 interactions with SARS-CoV-2-like viruses of pangolin origin. *EMBO J* **40**, e107786, doi:10.15252/embj.2021107786 (2021).
20 He, W. T. *et al.* Virome characterization of game animals in China reveals a spectrum of emerging pathogens. *Cell* **185**, 1117-1129 e1118, doi:10.1016/j.cell.2022.02.014 (2022).
21 Xiao, X., Newman, C., Buesching, C. D., Macdonald, D. W. & Zhou, Z. M. Animal sales from Wuhan wet markets immediately prior to the COVID-19 pandemic. *Sci Rep* **11**, 11898, doi:10.1038/s41598-021-91470-2 (2021).
22 Wang, Q. *et al.* Tracing the origins of SARS-CoV-2: lessons learned from the past. *Cell Res* **31**, 1139-1141, doi:10.1038/s41422-021-00575-w (2021).
23 Tong, Y. *et al.* The origins of viruses: discovery takes time, international resources, and cooperation. *Lancet* **398**, 1401-1402, doi:10.1016/S0140-6736(21)02180-2 (2021).
24 WHO-convened global study of origins of SARS-CoV-2: China Part. https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part (2021).
25 Tang, X. *et al.* On the origin and continuing evolution of SARS-CoV-2. *Natl Sci Rev* **7**, 2, doi: 10.1093/nsr/nwaa036 (2020).
26 Rambaut, A. *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* **5**, 1403-1407, doi:10.1038/s41564-020-0770-5 (2020).
27 Pekar, J. E. *et al.* The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science* **377**, 960-966, doi:10.1126/science.abp8337 (2022).
28 Worobey, M. *et al.* The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science* **377**, 951-959, doi:10.1126/science.abp8715 (2022).
29 Maxmen, A. Wuhan market was epicentre of pandemic's start, studies suggest. *Nature* **603**, 15-16, doi:10.1038/d41586-022-00584-8 (2022).
30 Li, H. *et al.* Human-animal interactions and bat coronavirus spillover potential among rural residents in Southern China. *Biosaf Health* **1**, 84-90, doi:10.1016/j.bsheal.2019.10.004 (2019).

Figure legends

Fig. 1. The distribution of the positive environmental samples in the Huanan Seafood Market.

A. As the place of the early cluster of COVID-19 patients, the Huanan Seafood Market is separated into East and West Zones with the Xinhua Road between them. To detect for the presence of SARS-CoV-2 RNA, reverse transcription, quantitative polymerase chain reaction (RT-qPCR) was performed. The locations of the positive samples were marked in the map of the market within orange, while the location of the samples that the live viruses were isolated from were labeled with red. The map also shows locations of stalls where domesticated wildlife products were sold. B. Timeline of environmental and animal samples collected within and around the Huanan Seafood Market. The information of confirmed patients up to December 31st 2019 was referenced from the Report of WHO-convened global study of origins of SARS-CoV-2.

Fig. 2. The SARS-CoV-2 virus isolation from environmental samples of the Huanan Seafood Market.

The electron micrographs of the SARS-CoV-2 viruses isolated from the environmental samples in the Huanan Seafood Market. To determine whether SARS-CoV-2 particles could be visualized from the cell supernatant and lysate, we used transmission electron microscopy (EM) to observe the culture supernatant and ultra-thin section cells based from both VeroE6 and Huh7.5 cells. The electron micrographs showed that virus particles were present in both the supernatant (A, B) and the cells (C, D). Negative-stained virus particles were generally spherical, pleomorphic and 60-140 nm in diameter. Spike protrusions were observed around the particles in a crown (corona) shape (A, B). In ultra-thin cultured cell sections, a group of virus particles can be seen outside the cell (C), and sheets of virus particles can also be observed inside the cells (D).

Fig. 3. Genomic and phylogenetic analyses of SARS-CoV-2 virus genomes from

the Huanan Seafood Market.

A. Sequence comparison of the full-length SARS-CoV-2 genomes in the environmental samples. B. Phylogenetic analysis of full-length SARS-CoV-2 genomes from the Huanan Seafood Market and representative strains from the early stage of the COVID-19 pandemic, showing that most environmental strains cluster together with the human strains in the B/L lineage, with Env_0020 in A/S lineage.

Fig. 4. Analysis of environmental samples in the Huanan Seafood Market.

A. Schematic illustration of the experimental design. All 73 SARS-CoV-2 positive samples were included for RNA-seq. A total of 60 RNA-seq libraries were successfully constructed. Additionally, RNA-seq libraries of 112 SARS-CoV-2 negative samples passed library quality control. The kranken2 was used for genus classification. The bowtie2 and sequences in the barcode of life data system was used for the classification of genus in the Mammalia class. B. Heatmap showing the reads distribution of the four domains (Bacteria, Eukaryota, Viruses and Archaea), the *Homo* genus, the *Mammalia* class and the SARS-CoV-2 species. SARS-CoV-2 PCR-positive or -negative were shown in the left panel. C. Positive ratio of illustrated genus in all tested samples. Top ranked genus within the *Mammalia* class were shown. D. Illustration of mammal genera in market using the threshold of 100 reads per millions. The samples were group by SARS-CoV-2 PCR results. The blue bar indicates the positive detected genera. E. Illustration of mammal genera distribution in samples with high viral load. The Env_0020, Env_0313, Env_0354 and Env_0126 were shown.

F. Distribution of the positively detected *Mammal* genera in the market. Samples in four areas where multiple SARS-CoV-2 PCR-positive samples were plotted. The distribution of top mammal genera in each area was shown.

Tables

Table 1. Information of the positive environmental samples from the Huanan Seafood Market[a].

Sample ID	Lab code	Sample type	Sampling date	Street No.	Vendor No.	PCR	CT	PCR target	NGS [c]
Env_0275	E61	Ground	2020/1/1	6	1-3	+	36.04	ORF1ab/N	Yes
Env_0101	A101	Surface of the door	2020/1/1	4	19	+	36.82	ORF1ab/N	Yes
Env_0014	A14	Surface of packing bag for hairtail	2020/1/1	4	26	+	36.42	ORF1ab/N	Yes
Env_0015	A15	Surface of the door	2020/1/1	4	26	+	35.51	ORF1ab/N	Yes
Env_0018	A18	Shoe covers and soles	2020/1/1	7	15-17	+	33.79	ORF1ab/N	Yes
Env_0002	A2	Ground	2020/1/1	7	15-17	+	35.07	ORF1ab/N	Yes
Env_0020	A20[f]	Gloves	2020/1/1	7	15-17	+	32.48	ORF1ab/N	Yes
Env_0033	A33	Garbage truck	2020/1/1	7	NA	+	34.46	ORF1ab/N	Yes
Env_0055	A55	Ground	2020/1/1	7	25	+	34.84	ORF1ab/N	Yes
Env_0061	A61	Ground	2020/1/1	7	20-22-24	+	32.04	ORF1ab/N	Yes
Env_0063	A63	Ground	2020/1/1	7	16-18	+	34.43	ORF1ab/N	Yes
Env_0087	A87	Surface of the door	2020/1/1	4	28	+	36.94	ORF1ab/N	Yes
Env_0088	A88	Ground	2020/1/1	4	28	+	36.69	ORF1ab/N	Yes
Env_0090	A90	Ground	2020/1/1	4	26	+	33.14	ORF1ab/N	Yes
Env_0096	A96	Ground	2020/1/1	4	24	+	33.97	ORF1ab/N	Yes
Env_0138	B17	Scale	2020/1/1	15	X44	+	34.16	ORF1ab/N	Yes
Env_0126	B5[b, f]	Ground	2020/1/1	5	6-8	+	29.32	ORF1ab/N	Yes
Env_0213	D32	Surface of a cart	2020/1/1	15	15	+	33.72	ORF1ab/N	Yes
Env_0262	E48	Blood on the ground in front of the door	2020/1/1	9	22	+	35.93	ORF1ab/N	Yes
Env_0221	E7	Styrofoam desk in front of the door	2020/1/1	2	5	+	36.44	ORF1ab/N	Yes
Env_0400	F100	Ground	2020/1/1	4	X6-X4	+	34.72	ORF1ab/N	Yes
Env_0313	F13[b, f]	Surface of the wall	2020/1/1	11	15	+	23.85	ORF1ab/N	Yes
Env_0333	F33	Roller shutter	2020/1/1	2	17	+	34.13	ORF1ab/N	Yes
Env_0346	F46	Ground	2020/1/1	2	24	+	31.8	ORF1ab/N	Yes
Env_0354	F54[b, f]	Ground	2020/1/1	2	14	+	25.8	ORF1ab/N	Yes
Env_0398	F98	Ground	2020/1/1	4	X6-X4	+	34	ORF1ab/N	Yes
Env_0509	G93	Sewage	2020/1/1	8	19-21-23	+	33.23	ORF1ab/N	Yes
Env_0552	Q37	Inner surface of the freezer	2020/1/12	8	25	-	NA [d]	ORF1ab/N	Yes
Env_0576	Q61	Cart1	2020/1/12	6	29-31-33	-	NA	ORF1ab/N	Yes
Env_0579	Q64	Cart2	2020/1/12	6	29-31-33	+	+ [d]	ORF1ab/N	Yes
Env_0583	Q68	Ground	2020/1/12	6	29-31-33	+	+	ORF1ab/N	Yes
Env_0584	Q69	Feather removal machine	2020/1/12	6	29-31-33	+	+	ORF1ab/N	Yes
Env_0585	Q70	Iron container in inner room	2020/1/12	6	29-31-33	-	NA	ORF1ab/N	Yes
Env_0660	1-27-33	Water drain	2020/1/27	6	29-33	+	36	ORF1ab	No
Env_0664	1-27-37	Water drain	2020/1/27	10	4	+	35	ORF1ab	No
Env_0682	1-29-4	Water drain	2020/1/29	4	5-7-9	+	36	ORF1ab	No
Env_0686	1-29-8	Water drain	2020/1/29	5	11	+	37	ORF1ab	Yes
Env_0717	8-25-D	Ground inside the stalls	2020/2/3	8	25	+	35.9	ORF1ab	Yes
Env_0719	8-25-M1	Ground inside the stalls	2020/2/3	8	25	+	36.5	ORF1ab	Yes
Env_0742	WS-1-1	West sewage well 1	2020/2/5	NA	NA	+	36.00	ORF1ab	Yes
Env_0743	WS-1-2	West sewage well 2	2020/2/5	NA	NA	+	38.00	ORF1ab	No
Env_0744	WS-1-3	West sewage well 3	2020/2/5	NA	NA	+	34.01	ORF1ab	No
Env_0745	WS-1-4	West sewage well 4	2020/2/5	NA	NA	+	38.00	ORF1ab	No
Env_0746	WS-2-1	West sewage well 1	2020/2/5	NA	NA	+	37.30	ORF1ab	Yes
Env_0747	WS-2-2	West sewage well 2	2020/2/5	NA	NA	+	37.65	ORF1ab	No
Env_0748	WS-2-3	West sewage well 2	2020/2/5	NA	NA	+	36.72	ORF1ab	Yes
Env_0750	WS-3-2	West sewage well 3	2020/2/5	NA	NA	+	37.60	ORF1ab	No
Env_0752	WS-4-2	West sewage well 4	2020/2/5	NA	NA	+	37.10	ORF1ab	Yes
Env_0762	zong-1	Water drain	2020/2/9	1	NA	+	34.94	ORF1ab	No
Env_0828	w-6-29-33	Water drain	2020/2/15	6	29-33	+	37.97	ORF1ab	Yes
Env_0829	w-zong-1	Water drain	2020/2/15	1	NA	+	36.54	ORF1ab	Yes

Env_0813	EWS-2#-2	East sewage well 2	2020/2/15	NA	NA	+	35.32	ORF1ab	Yes
Env_0815	EWS-3#-2	West sewage well 3	2020/2/15	NA	NA	+	36.05	ORF1ab	Yes
Env_0816	WWS-1#	West sewage well 1	2020/2/15	NA	NA	+	34.44	ORF1ab	Yes
Env_0817	WWS-1#-2	West sewage well 1	2020/2/15	NA	NA	+	33.63	ORF1ab	Yes
Env_0818	WWS-1#-3	West sewage well 1	2020/2/15	NA	NA	+	33.58	ORF1ab	Yes
Env_0820	WWS-2#-2	West sewage well 2	2020/2/15	NA	NA	+	37.44	ORF1ab	Yes
Env_0821	WWS-2#-3	West sewage well 2	2020/2/15	NA	NA	+	36.88	ORF1ab	Yes
Env_0830	wws-1#-0	West sewage well 1	2020/2/15	1	NA	+	33.75	ORF1ab	Yes
Env_0806	W-8-25-D1	Ground inside the stalls	2020/2/15	8	25	+	36.77	ORF1ab	Yes
Env_0807	W-8-25-D2	Ground inside the stalls	2020/2/15	8	25	+	33.91	ORF1ab	Yes
Env_0808	W-8-25-L	Container	2020/2/15	8	25	+	34.58	ORF1ab	Yes
Env_0809	W-8-25-L2	Container	2020/2/15	8	25	+	37.16	ORF1ab	Yes
Env_0838	C8	Wall inside the stalls	2020/2/20	5	stair1-2	+	32.59/-/37.34[e]	RdRp/N/E	No
Env_0862	SJ-D	Storehouse ground	2020/2/22	8	25	+	39.39/39.25/35.48	RdRp/N/E	No
Env_0863	SJ-CS	Storehouse weight scale	2020/2/22	8	25	+	40.21/40.1/36.37	RdRp/N/E	No
Env_0865	SJ-L3	Storehouse wire fence	2020/2/22	8	25	+	41.77/41.62/37.61	RdRp/N/E	No
Env_0867	RLC-4	Storehouse bag surfaces	2020/2/22	8	25	+	41.71/41.53/37.31	RdRp/N/E	No
Env_0868	RLC-3	Storehouse bag surfaces	2020/2/22	8	25	+	36.18/36.05/32.36	RdRp/N/E	Yes
Env_0601	WH-17	floor surface swab collected from Dongxihu Market	2020/1/22	NA	NA	+	33.90	ORF1ab	No
Env_0790	CSSQ-1-3	sewerage well in surrounding area	2020/2/9	NA	NA	+	37.23	ORF1ab	Yes
Env_0797	YCHC2-1-1	sewerage well in surrounding area	2020/2/9	NA	NA	+	36.42	ORF1ab	No
Env_0802	HXJC4-1-3	sewerage well in surrounding area	2020/2/9	NA	NA	+	36.05	ORF1ab	Yes

[a] Four positive samples were also included, that were not collected from the Huanan Seafood Market. One sample was a floor surface swab collected from Dongxihu Market in Wuhan on January 22th, 2020. and 3 collected from sewerage wells in the surrounding areas collected on February 9th, 2020.

[b] The live viruses were isolated from the three samples.

[c] "Yes", Samples with NGS analysis; "No", Samples without NGS analysis due to the limited RNA sample amount.

[d] The PCR results of Env_0552, Env_0576 and Env_0585 were negative. There were no CT values for these samples. The PCR results of Env_0579, Env_0583 and Env_0584 were positive, but the CT values for these samples were not recorded from the laboratory.

[e] The "-" indicates the negative result for the N gene target.

[f] The SARS-CoV-2 genomes were sequenced from the four environmental samples or their cell supernatants after virus isolation.

Table 2. The animal samples collected in the Huanan Seafood Market.

Species	Animal number	Sample number	RT-PCR positive number
Rabbit/Hares	52	104	0
Stray cat	27	80[a]	0
Snake	40	80	0
Hedgehog	16	67	0
Muntjac	6	18	0
Dog	7[b]	17	0
Badger	6	16	0
Bamboo rat	6	15	0
Rat[e]	10	12	0
Pig	NA[c]	6[d]	0
Chicken	5	5	0
Chinese giant salamander	3	5	0
Crocodile	2	4	0
Wild boar	2	4	0
Soft-shelled turtle	2	3	0
Weasel[e]	1	2	0
Fish	2	2	0
Sheep	1	1	0
Others	NA[f]	16	0
Total	188	457	0

[a] Six of the cats were from the Huanan Seafood Market. And the samples included faeces.

[b] Including one stray dog in the Huanan Seafood Market.

[c] Not applicable due to the processed pork.

[d] Collected from other markets.

[e] The weasel was not sold in the market, but caught alive in the Market.

[f] Not applicable due to the unrecognized “bai tiao” product as described in the Extended data methods.

Extended data methods

Sample collection

The Huanan Seafood Market (HSM) was closed in the early morning of January 1st 2020, and at the same time, China CDC began collecting environmental and animal samples. Staff from China CDC entered the market about 30 times before the market's final clean-up on March 2^{nd} 2020, with some stray animals sampled outside the market until March 30^{th}. Environmental samples in the HSM were collected to represent exhaustively as possible, from a wide diversity of surfaces, animals and products (Table 1 and 2) according to different sampling principles, as described in detail in the Joint Report of WHO-convened Global Study of Origins of SARS-CoV-2: China Part [1].

The principles and ranges of in-market sampling covered: (1) environmental samples from stalls related to early cases; (2) environmental samples from doors and floors of all stalls in the blocks where the early cases were located; (3) environmental samples in the East Zone of the market were collected according to blocks; (4) transport carts, trash cans and similar objects; (5) environmental samples from stalls that sold livestock, poultry, farmed wildlife (also called "domesticated wildlife" or "domesticated wildlife products" in this report); (6) samples of sewage and silt from drainage channels and sewerage wells; (7) stray cats, rats and other stray animals in the market; (8) animal products and other commodity samples kept in the cold storages and refrigerators in the market; (9) the market's ventilation and air-conditioning system; and (10) public toilets, public activity rooms and other places where people gathered in the market.

The investigators used full personal protective equipment during the sampling in the market. Commercial products of swabs and virus preservation solution were used for the sampling (Disposable Virus Sampling Tube, V5-S-25, Shen Zhen Zi Jian Biotechnology Co., Ltd., Shenzhen, China). For environmental samples, sampling swabs were applied to smear the floors, walls or surfaces of objects and then preserved them in virus preservation solution.

For animal samples, depending on the type of animal and whether it was alive or frozen, pharyngeal, anal, body surface and body cavity swabs or tissue samples were

collected for nucleic acid testing (NAT). Generally, for alive animal and frozen full bodies, three samples, including pharyngeal, anal and body surface swabs were collected for each animal individuals. And for animal bodies after “bai tiao” disposing (remaining parts of poultry or livestock after removal of hair and viscera), the body cavity swabs were collected.

Drain samples were collected by the use of virus sampling swabs to probe into the silt at the bottom of drainage channels in the market. Wastewater and silt samples were preserved in virus preservation solution. For the sewage well (for the drain water), a container was used to take a silt-water mixture from a location near the bottom of the well, and an appropriate amount of sample was collected by using virus sampling swabs and then preserved in virus preservation solution.

Nucleic acid extraction and SARS-CoV-2 real-time PCR assay

A virus nucleic acid extraction kit (Xi'an Tianlong) was used to extract viral nucleic acid from samples using an automated nucleic acid extraction instrument according to the manufacturer’s instructions. Real-time (RT) PCR was performed on extracted nucleic acid samples with a SARS-CoV-2 nucleic acid assay kit. The reagent brands include BioGerm (40/38, cycle number/cut-off value, the same as below), DAAN (45/40) and BGI (40/38).

Virus isolations

Virus isolations were performed in biosafety level (BSL)-3 laboratory in National Institute for Viral Diseases Control and Prevention, China CDC. Samples positive for SARS-CoV-2 were cultured in Vero E6 and Huh7.5 cells on January 11th, 2020. The cell lines were inoculated with positive samples and three blind passages were performed for each sample. The culture supernatant and cell pellet of each passage were harvested for RT PCR. The morphology of viral particles in the cell sections and the supernatant were firstly observed by transmission electron microscope (TEM) on January 22nd, 2020.

Metagenomic sequencing

Metagenomic sequencing was conducted at National Institute for Viral Disease Control

and Prevention, China CDC and Wuhan BGI. Nucleic acid was extracted using Qiagen's viral RNA microextraction kit and human nucleic acid was removed using an enrichment kit to improve the sensitivity of viral RNA detection. Extracted RNA was reverse transcribed into cDNA and segmented into 150-200 bp by enzyme digestion. After repair, fitting, purification, PCR amplification and purification, sample concentration was assayed by DNBSEQ-T7, and an average output of more than 200 million reads was obtained. Sequencing data were compared with those in a SARS-CoV-2 database to determine whether the samples contained coronavirus sequences. For the seven complete SARS-CoV-2 genome sequences, three sequences from environmental samples (Env_0020_seq01, Env_0313_seq02 and Env_0354_seq03) were obtained from DNBSEQ-T7, and four sequences from cell supernatants of Env_0313, Env_0354 and Env_0126 (Fig. 3) were obtained from NextSeq 550 platform. A few samples were re-sequenced using a multiplex PCR approach, including Env_0020_seq01, Env_0313_seq04, Env_0313_seq05, Env_0126_seq06, and Env_0354_seq07 (Extended Data Table 6 and 7). All raw data related to the genomes, including any partial genomes that were sequenced were fully reported and deposited to the public database (Extended Data Table 6 and 7).

Virus genome assembly and phylogenetic analysis

Raw reads were adaptor- and quality- trimmed with the Fastp (version 0.20.0) program. The clean reads were mapped to the SARS-CoV-2 reference genome (GenBank: NC_045512) using Bowtie2. The assembled genomes were merged and checked using Geneious (version 11.1.5) (https://www.geneious.com). The coverage and depth of genomes were calculated with SAMtools v1.10 based on SAM files from Bowtie2.

Reference genomes, IVDC-HB-01 (GISAID: EPI_ISL_402119) and Wuhan-Hu-1 (GenBank: NC_045512), were employed as a query. Multiple sequence alignment of the seven SARS-CoV-2 sequences obtained from this study and reference sequences were performed with Mafft (v7.450). Phylogenetic analyses were performed using RAxML v8.2.9 with 1000 bootstrap replicates, employing the GTR nucleotide substitution model and the Gamma distribution.

Bioinformatic analysis of the species abundances

The Kraken2 (version 2.1.2)[2] was used for species classification with the option '--confidence 0.1'. Sequences of all species in the Nucleotide (nt) database were used for generating the index. The bracken (version 2.5) was used for re-evaluating species abundance. The matrix of species was obtained by using the pavian algorithm[3]. ggplot2 package in R was used for plotting. Read counts of each genus were used for further analysis and plotting. Raw counts of four domains (Archaea, Viruses, Eukaryota, and Bacteria), SARS-CoV-2, *Homo* genus, and *Mammalia* class were shown by heatmap (4B). Two tail unpaired t-test was used for identification of differential genus between SARS-CoV-2 PCR-positive and -negative samples.

For the analysis of the mammalian genus characterization, the reference was generated using the sequence of mitochondrial cytochrome c oxidase subunit I (COI-5P) in the barcode of life data (BOLD) system[4-6]. RNA-seq samples were mapped to the reference sequences by the bowtie2[7] algorithm with the default settings. Read counts of each genus were calculated by the samtools[8]. Read counts over 20 were used as cut-off for the identification of positively enriched genus. Fisher's exact test was used for comparing the differential genus in the Mammalia class between SARS-CoV-2 PCR-positive and -negative samples.

Ethics

The sample collection was determined by China CDC to be part of the emergency responses to the pneumonia of unknown etiology (PUE) and therefore exempt from institutional review board assessment.

Methods References

1 WHO-convened global study of origins of SARS-CoV-2: China Part. *https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part* (2021).

2 Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol* **20**, 257, doi:10.1186/s13059-019-1891-0 (2019).

3 Breitwieser, F. P. & Salzberg, S. L. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics* **36**, 1303-1304, doi:10.1093/bioinformatics/btz715 (2020).

4 Valentini, A., Pompanon, F. & Taberlet, P. DNA barcoding for ecologists. *Trends Ecol Evol* **24**, 110-117, doi:10.1016/j.tree.2008.09.011 (2009).

5 Hebert, P. D., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. Identification of birds through DNA barcodes. *PLoS Biol* **2**, e312, doi:10.1371/journal.pbio.0020312 (2004).

6 Ratnasingham, S. & Hebert, P. D. bold: The barcode of Life Data System (http://www.barcodinglife.org). *Mol Ecol Notes* **7**, 355-364, doi:10.1111/j.1471-8286.2007.01678.x (2007).

7 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

8 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

Extended Data Figure Legends

Extended Data Fig. 1. The overground drainage pathway in the Huanan Seafood Market and environmental sample collection.

The wastewater in the overground drainage was lead into the underground drainage inside the market and then flow into the wells on the edge of the market. And we did a spot-check sampling across all the overground drainages. To detect for the presence of SARS-CoV-2 RNA, reverse transcription and quantitative polymerase chain reaction (RT-qPCR) were performed. The locations of the positive samples were marked in the map of the market within yellow.

Extended Data Fig. 2. Positive environmental samples associated with different products in the Huanan Seafood Market.

Dots represent the percentage of positive environmental samples associated with each product. Bars represent 95% confidence intervals for the binomials in the text above. Note that the confidence interval (CI) for some products (e.g. vegetables, farmed wildlife) have broad error bars that are likely due to the low number of vendors for these categories in the market. Nine of the 10 vendors selling farmed wildlife have been sampled.

Extended Data Fig. 3. Illustration of mammal genera distribution in samples of concerns. Illustration of mammal genera distribution in samples of concerns. Samples related to the blood spot and the de-feather machine (Env_0262 and Env_0584) and samples enriched with genera related to wildlife (Env_0576, Env_0807, Env_0809, and Env_0585) were plotted. Animal genera identified by the BOLD method were shown in the left panel, while mammal genera identified by the kraken2 method were shown in the right panel.

Extended Data Fig. 4. Distribution of the positively detected Mammal genera in the market. The distribution of SARS-CoV-2 and potential host were plotted by yellow and blue dots, respectively. The density of the distribution of potential host was shown in red, while the SARS-CoV-2 by green.

Supplementary information guide for the SI files

Supplementary Table 1. The information of the 1380 samples collected from or around the Huanan Seafood Market (HSM). Sample ID: Unique ID for each sample. Lab code: Temporary IDs used for sampling and laboratory testing. Sampling location: The relative location of the sampling site to the HSM, or the relationship of the site to the SHM, whether it is inside the HSM (east/west), near the HSM, or other. Type of vendor sold product including aquatic, seafood, poultry, livestock, wildlife, vegetable, and cold-chain product: Yes, means that the vendor where the sample was collected sells this type of product, and No means that the vendor where the sample was collected does not sell this type of product. These pieces of information together indicate the scope of business of the vendor where the sample was collected.

Supplementary Table 2. Distribution of different genus in each sample.

Supplementary Table 3. Classification of genus within the Mammalia class using the barcode of life data (BOLD) system.

Supplementary Table 4. The differential of genus between SARS-CoV-2 PCR-positive and -negative environmental samples.

Supplementary Table 5. Risk of co-existence with SARS-CoV-2 positive samples.

Extended Data Tables

Extended Data Table 1. Overview of environmental sample sampling and testing in the Huanan Seafood Market.

	Number of samples	Number of positive samples by RT-PCR	Number of isolated viruses
Huanan Seafood Market	718	40	3
Warehouses related to the Huanan Seafood Market[a]	14	5	
Other markets in Wuhan and Huanggang[b]	30	1	
Drainage system in the Huanan Seafood Market	110	24	
Sewerage wells in surrounding areas	51	3	
Total	923	73	3

[a] The warehouses related to the Huanan Seafood Market were located out of the market.

[b] The one positive sample outside HSM was collected from Dongxihu Market in Wuhan. More information was provided in Extended Data Table 4.

Extended Data Table 2. The collection logic of the environment samples.

No.	Time	Objective	Sample time	Amount	Sum
1	1,Jan	(1) Environmental samples from stalls related to early cases; (2) Environmental samples from doors and floors of all stalls in the blocks where the early cases were located; (3) Environmental samples in the east wing of the market were collected according to blocks; (4) Transport carts, trash cans and similar objects.	1,Jan	515	515
2	12,Jan	Environmental samples from stalls that sold livestock, poultry, farmed wildlife (also called domesticated wildlife).	12,Jan	70	70
3	22,Jan	Environmental samples from other markets in Wuhan	22,Jan	30	30
4	23,Jan-19,Feb	The outdoor environmental samples from stalls that sold livestock, poultry, farmed wildlife.	23,Jan	23	52
			25,Jan	2	
			3,Feb	16	
			9,Feb	5	
			15,Feb	4	
			19,Feb	2	
5	27,Jan-15,Feb	Samples of sewage and silt from drainage channels and sewerage wells in the market.	27,Jan	38	94
			29,Jan	26	
			9,Feb	9	
			15,Feb	21	
6	5,Feb-9,Feb	Samples of sewage and silt from city sewerage wells around the market.	5,Feb	32	71
			9,Feb	39	
7	20,Feb-2,Mar	(1) Cold storages and refrigerators from stalls that sold livestock, poultry, farmed wildlife in the market; (2) The market's ventilation and air-conditioning system; (3) Public toilets, public activity rooms and other places where people gathered in the market.	20,Feb	27	91
			22,Feb	12	
			23,Feb	1	
			25,Feb	2	
			29,Feb	15	
			2,Mar	34	
Total				923	

Extended Data Table 3. The collection logic of the animal samples.

No.[a]	Time	Objectives	Sample time	Amount	Sum
8	22,Jan	Animal products in other markets.	22,Jan	6	6
9	25,Jan-10,Mar	Animal products and other commodity samples kept in the cold storages and refrigerators in the market.	25,Jan	55	306
			20,Feb	23	
			21,Feb	36	
			23,Feb	5	
			25,Feb	47	
			2,Mar	75	
			10,Mar	65	
10	27,Jan-1,Mar	Live animals captured around the market.	27,Jan	5	17
			5,Feb	3	
			9,Feb	2	
			15,Feb	3	
			29,Feb	2	
			1,Mar	2	
11	18,Jan-30,Mar	Stray cats, mice, cat feces and other stray animals (one dog and one weasel in the market).	18,Jan	1	96
			27,Jan	12	
			28,Jan	8	
			29,Jan	21	
			5,Feb	10	
			15,Feb	2	
			23,Feb	2	
			14,Mar	2	
			20,Mar	2	
			22,Mar	4	
			30,Mar	32	
12	19,Feb-23,Feb	Animal products and other commodity samples kept in the cold storages.	19,Feb	28	32
			23,Feb	4	
Total				457	

[a] The number follows the upper Table for environment samples.

Extended Data Table 4. The information of the sampling in other markets.

District	Number of environment samples [a]	Number of positive environment samples by RT-PCR	Number of animal samples [b]	Number of positive animal samples by RT-PCR
Jiang'han district	7	0	2	0
Jiang'an district	8	0	2	0
Donxihu district	7	1	1	0
Huanggang city	8	0	1	0
Total	30	1	6	0

[a] Swab sample collected from the floor, wall or chopping board.

[b] The heart, liver and large intestine tissues from pigs.

Extended Data Table 5. Twenty-one venders of nucleic acid testing (NAT) positive in the Huanan Seafood Market.

Vendors No.	Location	Product types [a]						
		Cold-chain products	Aquatic products	Seafood products	Poultry	Livestock	Wildlife products	Vegetables
1	West	no	no	no	yes	no	no	no
2	West	yes	yes	yes	no	no	no	no
3	West	yes	yes	no	yes	yes	yes	no
4	East	yes	no	no	yes	yes	no	no
5	West	no	no	no	no	no	no	no
6	West	no	yes	no	yes	yes	no	no
7	West	yes	no	no	yes	no	no	no
8	West	yes	yes	yes	yes	no	no	no
9	West	yes	yes	yes	no	no	no	no
10	West	yes	yes	yes	yes	yes	no	no
11	West	yes	yes	no	no	no	no	no
12	West	yes	yes	yes	no	no	no	no
13	West	yes	yes	no	no	no	no	no
14	West	yes	yes	no	no	no	no	no
15	West	yes	yes	no	no	no	no	no
16	West	yes	yes	no	no	no	no	no
17	West	no	no	no	no	no	no	no
18	West	yes	no	no	yes	yes	no	no
19	West	no	no	no	no	no	no	yes
20	West	yes	no	no	no	no	no	yes
21	East	yes	yes	yes	no	no	no	no
Sum of NAT positive vendors		16	13	6	8	5	1	2
Vendors sampled in the study selling such products		87	73	56	37	36	9	8

[a] "yes" indicates product sold by vendors; "no" indicates product not sold by vendors.

761 **Extended Data Table 6. Information of the high-throughput sequencing data of**
762 **environmental samples from Huanan Seafood Market[a].**

No.	Sample ID	Lab Code	GISAID_Accession[b]	SARS-CoV-2 qPCR result	Sample type	Street No.[c]	Vendor No.
1	Env_0002	A2	EPI_ISL_13052326	Positive	Ground	7	15-17
2	Env_0014	A14	EPI_ISL_13052323	Positive	Surface of packing bag for hairtail	4	26
3	Env_0015	A15	EPI_ISL_13052324	Positive	Surface of the door	4	26
4	Env_0018	A18	EPI_ISL_13052325	Positive	Shoe covers and soles	7	15-17
5	Env_0020	A20	EPI_ISL_13052327	Positive	Gloves	7	15-17
6	Env_0033	A33	EPI_ISL_13052328	Positive	Garbage cart	7	NA
7	Env_0055	A55	EPI_ISL_13052329	Positive	Ground	7	25
8	Env_0061	A61	EPI_ISL_13052330	Positive	Ground	7	20-22-24
9	Env_0063	A63	EPI_ISL_13052331	Positive	Ground	7	16-18
10	Env_0087	A87	EPI_ISL_13052332	Positive	Surface of the door	4	28
11	Env_0088	A88	EPI_ISL_13052333	Positive	Ground	4	28
12	Env_0090	A90	EPI_ISL_13052334	Positive	Ground	4	26
13	Env_0096	A96	EPI_ISL_13052335	Positive	Ground	4	24
14	Env_0101	A101	EPI_ISL_13052322	Positive	Surface of the door	4	19
15	Env_0126	B5	EPI_ISL_13052337	Positive	Ground	5	6-8
16	Env_0138	B17	EPI_ISL_13052336	Positive	Scale	15	X44
17	Env_0213	D32	EPI_ISL_13052338	Positive	Surface of a cart	15	15
18	Env_0221	E7	EPI_ISL_13052341	Positive	Styrofoam desk in front of the door	2	5
19	Env_0262	E48	EPI_ISL_13052339	Positive	Blood on the ground in front of the door	9	22
20	Env_0275	E61	EPI_ISL_13052340	Positive	Ground	6	1-3
21	Env_0313	F13	EPI_ISL_13052343	Positive	Surface of the wall	11	15
22	Env_0346	F46	EPI_ISL_13052344	Positive	Ground	2	24
23	Env_0354	F54	EPI_ISL_13052345	Positive	Ground	2	14
24	Env_0398	F98	EPI_ISL_13052346	Positive	Ground	4	X6-X4
25	Env_0400	F100	EPI_ISL_13052342	Positive	Ground	4	X6-X4
26	Env_0552	Q37	EPI_ISL_13052297	Positive	Inner surface of the freezer	8	25
27	Env_0576	Q61	EPI_ISL_13052298	Positive	Cart1	6	29-31-33
28	Env_0579	Q64	EPI_ISL_13052299	Positive	Cart2	6	29-31-33
29	Env_0583	Q68	EPI_ISL_13052300	Positive	Ground	6	29-31-33
30	Env_0584	Q69	EPI_ISL_13052301	Positive	Feather removal machine	6	29-31-33
31	Env_0585	Q70	EPI_ISL_13052302	Positive	Iron container in inner room	6	29-31-33
32	Env_0660	1-27-33	EPI_ISL_17064533	Positive	Water drain	6	29-33
33	Env_0664	1-27-37	EPI_ISL_17064534	Positive	Water drain	10	4
34	Env_0682	1-29-4	EPI_ISL_17064535	Positive	Water drain	4	5-7-9

35	Env_0686	1-29-8	EPI_ISL_13052303	Positive	Water drain	5	11
36	Env_0717	8-25-D	EPI_ISL_17064536	Positive	Ground inside the stalls	8	25
37	Env_0719	8-25-M1	EPI_ISL_13052304	Positive	Ground inside the stalls	8	25
38	Env_0742	WS-1-1	EPI_ISL_13052305	Positive	Sewage well	NA	NA
39	Env_0746	WS-2-1	EPI_ISL_13052306	Positive	Sewage well	NA	NA
40	Env_0748	WS-2-3	EPI_ISL_13052307	Positive	Sewage well	NA	NA
41	Env_0752	WS-4-2	EPI_ISL_17064537	Positive	Sewage well	NA	NA
42	Env_0790	CSSQ-1-3	EPI_ISL_13052308	Positive	Sewage well	NA	NA
43	Env_0802	HXJC4-1-3	EPI_ISL_17064538	Positive	Sewage well	NA	NA
44	Env_0806	W-8-25-D1	EPI_ISL_13052309	Positive	Ground inside the stalls	8	25
45	Env_0807	W-8-25-D2	EPI_ISL_13052310	Positive	Ground inside the stalls	8	25
46	Env_0808	W-8-25-L	EPI_ISL_13052311	Positive	Container	8	25
47	Env_0809	W-8-25-L2	EPI_ISL_13052312	Positive	Container	8	25
48	Env_0813	EWS-2#-2	EPI_ISL_13052313	Positive	Sewage well	NA	NA
49	Env_0815	EWS-3#-2	EPI_ISL_13052314	Positive	Sewage well	NA	NA
50	Env_0816	WWS-1#	EPI_ISL_13052315	Positive	Sewage well	NA	NA
51	Env_0817	WWS-1#-2	EPI_ISL_13052316	Positive	Sewage well	NA	NA
52	Env_0818	WWS-1#-3	EPI_ISL_13052317	Positive	Sewage well	NA	NA
53	Env_0820	WWS-2#-2	EPI_ISL_17064539	Positive	Sewage well	NA	NA
54	Env_0821	WWS-2#-3	EPI_ISL_13052318	Positive	Sewage well	NA	NA
55	Env_0828	w-6-29-33	EPI_ISL_13052319	Positive	Water drain	6	29-33
56	Env_0829	w-zong-1	EPI_ISL_13052320	Positive	Water drain	1	NA
57	Env_0830	wws-1#-0	EPI_ISL_13052321	Positive	Sewage well	1	NA
58	Env_0862	SJ-D	EPI_ISL_17064540	Positive	Storehouse ground	8	25
59	Env_0865	SJ-L3	EPI_ISL_17064541	Positive	Storehouse wire fence	8	25
60	Env_0868	RLC-3	EPI_ISL_17064542	Positive	Storehouse bag surfaces	8	25
61	Env_0516	HJ200001-20200112-1	EPI_ISL_17064543	Negative	Table top	9	34-36
62	Env_0517	HJ200002-20200112-1	EPI_ISL_17064544	Negative	Scale	9	34-36
63	Env_0518	HJ200003-20200112-1	EPI_ISL_17064545	Negative	Container	9	34-36
64	Env_0519	HJ200004-20200112-1	EPI_ISL_17064546	Negative	Container	9	34-36
65	Env_0520	HJ200005-20200112-1	EPI_ISL_17064547	Negative	Basket	9	34-36
66	Env_0521	HJ200006-20200112-1	EPI_ISL_17064548	Negative	Feather removal machine	9	34-36

67	Env_0522	HJ200007-20200112-1	EPI_ISL_17064549	Negative	Table	9	34-36
68	Env_0523	HJ200008-20200112-1	EPI_ISL_17064550	Negative	Inner surface of refrigerator	9	34-36
69	Env_0524	HJ200009-20200112-1	EPI_ISL_17064551	Negative	Inside of feather removal machine	9	34-36
70	Env_0525	HJ200010-20200112-1	EPI_ISL_17064552	Negative	Outside surface of refrigerator	9	34-36
71	Env_0526	HJ200011-20200112-1	EPI_ISL_17064553	Negative	Wall	9	38
72	Env_0527	HJ200012-20200112-1	EPI_ISL_17064554	Negative	White basket	9	38
73	Env_0528	HJ200013-20200112-1	EPI_ISL_17064555	Negative	Chopping block	9	38
74	Env_0529	HJ200014-20200112-1	EPI_ISL_17064556	Negative	Container	9	38
75	Env_0530	HJ200015-20200112-1	EPI_ISL_17064557	Negative	Container	9	38
76	Env_0531	HJ200016-20200112-1	EPI_ISL_17064558	Negative	Red basket	9	38
77	Env_0532	HJ200017-20200112-1	EPI_ISL_17064559	Negative	Scale	9	38
78	Env_0533	HJ200018-20200112-1	EPI_ISL_17064560	Negative	Bucket	9	38
79	Env_0534	HJ200019-20200112-1	EPI_ISL_17064561	Negative	Bucket	9	38
80	Env_0535	HJ200020-20200112-1	EPI_ISL_17064562	Negative	Wall	9	38
81	Env_0536	HJ200021-20200112-1	EPI_ISL_17064563	Negative	Ground	9	35-37
82	Env_0537	HJ200022-20200112-1	EPI_ISL_17064564	Negative	Outside surface of freezer	9	35-37
83	Env_0538	HJ200023-20200112-1	EPI_ISL_17064565	Negative	Wall	9	35-37
84	Env_0539	HJ200024-20200112-1	EPI_ISL_17064566	Negative	Blue mop	9	35-37
85	Env_0540	HJ200025-20200112-1	EPI_ISL_17064567	Negative	Tongs	9	35-37
86	Env_0541	HJ200026-20200112-1	EPI_ISL_17064568	Negative	Ground	9	35-37
87	Env_0542	HJ200027-20200112-1	EPI_ISL_17064569	Negative	Bucket	9	35-37
88	Env_0543	HJ200028-20200112-1	EPI_ISL_17064570	Negative	Basket	9	35-37

89	Env_0544	HJ200029-20200112-1	EPI_ISL_17064571	Negative	Surface of the door	9	35-37
90	Env_0545	HJ200030-20200112-1	EPI_ISL_17064572	Negative	Container in inner room	9	35-37
91	Env_0546	HJ200031-20200112-1	EPI_ISL_17064573	Negative	Basket	8	25
92	Env_0547	HJ200032-20200112-1	EPI_ISL_17064574	Negative	Foam box	8	25
93	Env_0548	HJ200033-20200112-1	EPI_ISL_17064575	Negative	Wall	8	25
94	Env_0549	HJ200034-20200112-1	EPI_ISL_17064576	Negative	Container	8	25
95	Env_0550	HJ200035-20200112-1	EPI_ISL_17064577	Negative	Outside surface of freezer	8	25
96	Env_0551	HJ200036-20200112-1	EPI_ISL_17064578	Negative	Iron container	8	25
97	Env_0553	HJ200038-20200112-1	EPI_ISL_17064579	Negative	Knief	8	25
98	Env_0554	HJ200039-20200112-1	EPI_ISL_17064580	Negative	Scale	8	25
99	Env_0555	HJ200040-20200112-1	EPI_ISL_17064581	Negative	Ground	8	25
100	Env_0556	HJ200041-20200112-1	EPI_ISL_17064582	Negative	Feather removal machine	8	36-38
101	Env_0557	HJ200042-20200112-1	EPI_ISL_17064583	Negative	Basket	8	36-38
102	Env_0558	HJ200043-20200112-1	EPI_ISL_17064584	Negative	Mop	8	36-38
103	Env_0559	HJ200044-20200112-1	EPI_ISL_17064585	Negative	Sink	8	36-38
104	Env_0560	HJ200045-20200112-1	EPI_ISL_17064586	Negative	Iron container	8	36-38
105	Env_0561	HJ200046-20200112-1	EPI_ISL_17064587	Negative	Basket	8	36-38
106	Env_0562	HJ200047-20200112-1	EPI_ISL_17064588	Negative	Bucket	8	36-38
107	Env_0563	HJ200048-20200112-1	EPI_ISL_17064589	Negative	Inner surface of freezer	8	36-38
108	Env_0564	HJ200049-20200112-1	EPI_ISL_17064590	Negative	Dustpan	8	36-38
109	Env_0565	HJ200050-20200112-1	EPI_ISL_17064591	Negative	Table top	8	36-38
110	Env_0566	HJ200051-20200112-1	EPI_ISL_17064592	Negative	Ground	8	37

111	Env_0567	HJ200052-20200112-1	EPI_ISL_17064593	Negative	Basket	8	37
112	Env_0568	HJ200053-20200112-1	EPI_ISL_17064594	Negative	Outside surface of refrigerator	8	37
113	Env_0569	HJ200054-20200112-1	EPI_ISL_17064595	Negative	Basket	8	37
114	Env_0570	HJ200055-20200112-1	EPI_ISL_17064596	Negative	Bucket	8	37
115	Env_0571	HJ200056-20200112-1	EPI_ISL_17064597	Negative	Inner surface of refrigerator	8	37
116	Env_0572	HJ200057-20200112-1	EPI_ISL_17064598	Negative	Table top	8	37
117	Env_0573	HJ200058-20200112-1	EPI_ISL_17064599	Negative	Inner surface of refrigerator	8	37
118	Env_0574	HJ200059-20200112-1	EPI_ISL_17064600	Negative	Wall	8	37
119	Env_0575	HJ200060-20200112-1	EPI_ISL_17064601	Negative	Iron shelf	8	37
120	Env_0577	HJ200062-20200112-1	EPI_ISL_17064602	Negative	Container	6	29-31-33
121	Env_0578	HJ200063-20200112-1	EPI_ISL_17064603	Negative	Basket	6	29-31-33
122	Env_0580	HJ200065-20200112-1	EPI_ISL_17064604	Negative	White box	6	29-31-33
123	Env_0581	HJ200066-20200112-1	EPI_ISL_17064605	Negative	Iron container	6	29-31-33
124	Env_0582	HJ200067-20200112-1	EPI_ISL_17064606	Negative	Bucket	6	29-31-33
125	Env_0617	10-31-In2	EPI_ISL_17064607	Negative	Environmental swab	10	31
126	Env_0619	10-31-abv2	EPI_ISL_17064608	Negative	Environmental swab	10	31
127	Env_0620	06-29-abv1	EPI_ISL_17064609	Negative	Environmental swab	6	29
128	Env_0642	1-27-11	EPI_ISL_17064610	Negative	Water drain	3	13
129	Env_0643	1-27-12	EPI_ISL_17064611	Negative	Water drain	1	2
130	Env_0657	1-27-28	EPI_ISL_17064612	Negative	Water drain	7	34
131	Env_0674	1-27-52	EPI_ISL_17064613	Negative	Water drain	15	22
132	Env_0697	1-29-19	EPI_ISL_17064614	Negative	Water drain	9	35
133	Env_0701	1-29-23	EPI_ISL_17064615	Negative	Ground	11	9
134	Env_0705	7-26-BX	EPI_ISL_17064616	Negative	Inner surface of refrigerator	7	26
135	Env_0707	7-26-PSW	EPI_ISL_17064617	Negative	Drainage outlet	7	26
136	Env_0708	7-26-PSN	EPI_ISL_17064618	Negative	Drainage outlet	7	26
137	Env_0712	7-26-DH	EPI_ISL_17064619	Negative	Ground	7	26
138	Env_0714	8-25-BX	EPI_ISL_17064620	Negative	Inner surface of refrigerator	8	25

139	Env_0715	8-25-D	EPI_ISL_17064621	Negative	Knief	8	25
140	Env_0716	8-25-CK	EPI_ISL_17064622	Negative	Sewage	8	25
141	Env_0720	8-25-M2	EPI_ISL_17064623	Negative	Ground	8	25
142	Env_0749	WS-3-1	EPI_ISL_17064624	Negative	Sewage well	NA	NA
143	Env_0751	WS-4-1	EPI_ISL_17064625	Negative	Sewage well	NA	NA
144	Env_0753	8-25-D1	EPI_ISL_17064626	Negative	Ground	8	25
145	Env_0754	8-25-D2	EPI_ISL_17064627	Negative	Ground	8	25
146	Env_0756	8-25-Long	EPI_ISL_17064628	Negative	Container	8	25
147	Env_0759	Outside-5	EPI_ISL_17064629	Negative	Water drain	5	NA
148	Env_0768	HXJMHL-1-2	EPI_ISL_17064630	Negative	Sewage well	NA	NA
149	Env_0770	JXRJ-1-1	EPI_ISL_17064631	Negative	Sewage well	NA	NA
150	Env_0788	CSSQ-1-1	EPI_ISL_17064632	Negative	Sewage well	NA	NA
151	Env_0789	CSSQ-1-2	EPI_ISL_17064633	Negative	Sewage well	NA	NA
152	Env_0798	YCHC2-1-2	EPI_ISL_17064634	Negative	Sewage well	NA	NA
153	Env_0799	YCHC2-1-3	EPI_ISL_17064635	Negative	Sewage well	NA	NA
154	Env_0800	HXJC4-1-1	EPI_ISL_17064636	Negative	Sewage well	NA	NA
155	Env_0801	HXJC4-1-2	EPI_ISL_17064637	Negative	Sewage well	NA	NA
156	Env_0839	C9	EPI_ISL_17064638	Negative	Ground	5	NA
157	Env_0842	C12	EPI_ISL_17064639	Negative	Wall	5	NA
158	Env_0847	C17	EPI_ISL_17064640	Negative	Surface of container	6	34
159	Env_0856	E-10-29-2	EPI_ISL_17064641	Negative	Container	10	29
160	Env_0858	E-A-7-1	EPI_ISL_17064642	Negative	Container	Accessory Street	7
161	Env_0873	CCDC-454	EPI_ISL_17064643	Negative	Container	9	34
162	Env_0875	629-1-L1	EPI_ISL_17064644	Negative	Container	6	29
163	Env_0877	629-3-C	EPI_ISL_17064645	Negative	Cart	6	29
164	Env_0879	629-5-L4	EPI_ISL_17064646	Negative	Container	6	29
165	Env_0882	629-8-L7	EPI_ISL_17064647	Negative	Container	6	29
166	Env_0885	629-11-L	EPI_ISL_17064648	Negative	Container	6	29
167	Env_0887	629-13-L	EPI_ISL_17064649	Negative	Container	6	29
168	Env_0889	629-L-1	EPI_ISL_17064650	Negative	Container	6	29
169	Env_0893	CCDC-744	EPI_ISL_17064651	Negative	Ground	6	29
170	Env_0895	CCDC-746	EPI_ISL_17064652	Negative	Container	6	29
171	Env_0907	CCDC-758	EPI_ISL_17064653	Negative	Container	8	25
172	Env_0909	CCDC-760	EPI_ISL_17064654	Negative	Ground	8	25

[a] All the raw sequencing data have been uploaded onto the GISAID (China CDC Weekly, 2021, DOI: 10.46234/ccdcw2021.255). The raw sequence data reported in this paper have also been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics, 2021, DOI: 10.1016/j.gpb.2021.08.001) in National Genomics Data Center (Nucleic Acids Res, 2022, DOI: 10.1093/nar/gkab951), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences that are publicly accessible at https://ngdc.cncb.ac.cn/gsa (GSA: CRA010170). Raw sequence data was deposited into NCBI BioProject under accession number PRJNA948658

(http://www.ncbi.nlm.nih.gov/bioproject/948658) and in China National Microbiology Data Center (NMDC) with accession numbers NMDC10018366 (https://nmdc.cn/resource/genomics/sample/detail/NMDC10018366).

[b] Three original samples (Env_0020, Env_0313, and Env_0354) were sequenced twice, for metagenomic analysis (as shown in this table) and viral whole-genome assembly (raw data was shown in Extend Data Table 7), respectively.

[c] The street or vendor information of some samples was not applicable, such as transport carts, trash cans, sewage and silt from drainage channels and sewerage wells.

Extended Data Table 7. Information of the three SARS-CoV-2 sequences from environmental samples and the four sequence from cell supernatants[a].

Sample ID	Lab code	Name	GISAID_Accession		Collection date [c]	Length	Mapped reads	Depths(mean±SD)	Passage details [d]
			Genome	Raw data [b]					
Env_0020_seq01	A20	hCoV-19/env/Wuhan/IVDC-HBA20/2020	EPI_ISL_10497477	EPI_ISL_17064655	2020/1/1	29368	897978	2902.8±1924.9	Original
Env_0313_seq02	F13	hCoV-19/env/Wuhan/IVDC-HBF13/2020	EPI_ISL_408511	EPI_ISL_17064656	2020/1/1	28557	187396	611.5±275.3	Original
Env_0354_seq03	F54	hCoV-19/env/Wuhan/IVDC-HBF54/2020	EPI_ISL_408512	EPI_ISL_17064657	2020/1/1	29820	52348	170.8±78.6	Original
Env_0313_seq04	F13-20	hCoV-19/env/Wuhan/IVDC-HBF13-20/2020	EPI_ISL_408514	EPI_ISL_17064658	2020/1/20	29820	225356	1117.6±934.6	P1
Env_0313_seq05	F13-21	hCoV-19/env/Wuhan/IVDC-HBF13-21/2020	EPI_ISL_408515	EPI_ISL_17064659	2020/1/21	29820	301488	1474.7±1362.8	P1
Env_0126_seq06	B5-P3	hCoV-19/env/Wuhan/IVDC-HBB5-P3/2020	EPI_ISL_10497479	EPI_ISL_17064660	2020/2/29	29865	1506094	3708.0±1150.8	P3
Env_0354_seq07	F54-P3	hCoV-19/env/Wuhan/IVDC-HBF54-P3/2020	EPI_ISL_10497481	EPI_ISL_17064661	2020/2/29	29864	1442240	3553.0±1143.0	P3

[a] All the raw sequencing data (different from the Extended Data Table 6) and 7 genome sequences have been uploaded onto the GISAID (China CDC Weekly, 2021, DOI: 10.46234/ccdcw2021.255). The raw sequence data reported in this paper have also been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics, 2021, DOI: 10.1016/j.gpb.2021.08.001) in National Genomics Data Center (Nucleic Acids Res, 2022, DOI: 10.1093/nar/gkab951), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences that are publicly accessible at https://ngdc.cncb.ac.cn/gsa (GSA: CRA010170). The viral genomes reported in this paper have been deposited in the GenBase in National Genomics Data Center (Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation under accession number C_AA002295.1 to C_AA002301.1 that is publicly accessible at nttps://ngdc.cnc.ac.cn/genbase. Raw sequence data was deposited into NCBI BioProject under accession number PRJNA948658 (http://www.ncbi.nlm.nih.gov/bioproject/948658) and in China National Microbiology Data Center (NMDC) with accession numbers NMDC10018366

(https://nmdc.cn/resource/genomics/sample/detail/NMDC10018366).

[b] Three original samples (Env_0020, Env_0313, and Env_0354) were sequenced twice, for metagenomic analysis (raw data was shown in Extend Data Table 6) and viral whole-genome assembly (as shown in this table: Env_0020_seq01, Env_0313_seq02, and Env_0354_seq03), respectively.

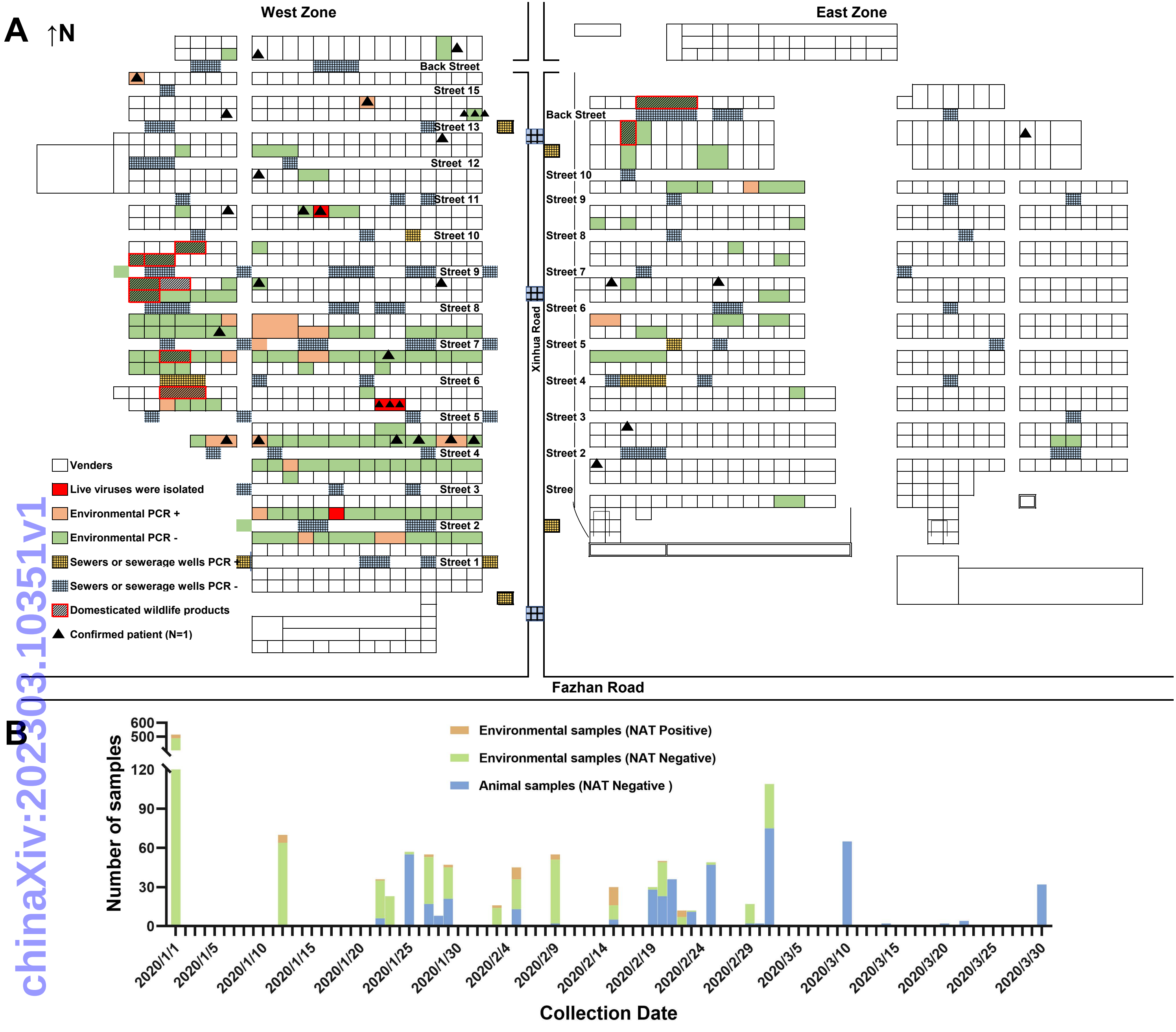
[c] The collection dates for the three SARS-CoV-2 sequences (Env_0020_seq01, Env_0313_seq02, and Env_0354_seq03) were the original environmental sample collection dates. And the collection dates for the other four sequences (Env_0313_seq04, Env_0313_seq05, Env_0126_seq06, Env_0354_seq07) were the date for the sequencing of the cell supernatants.

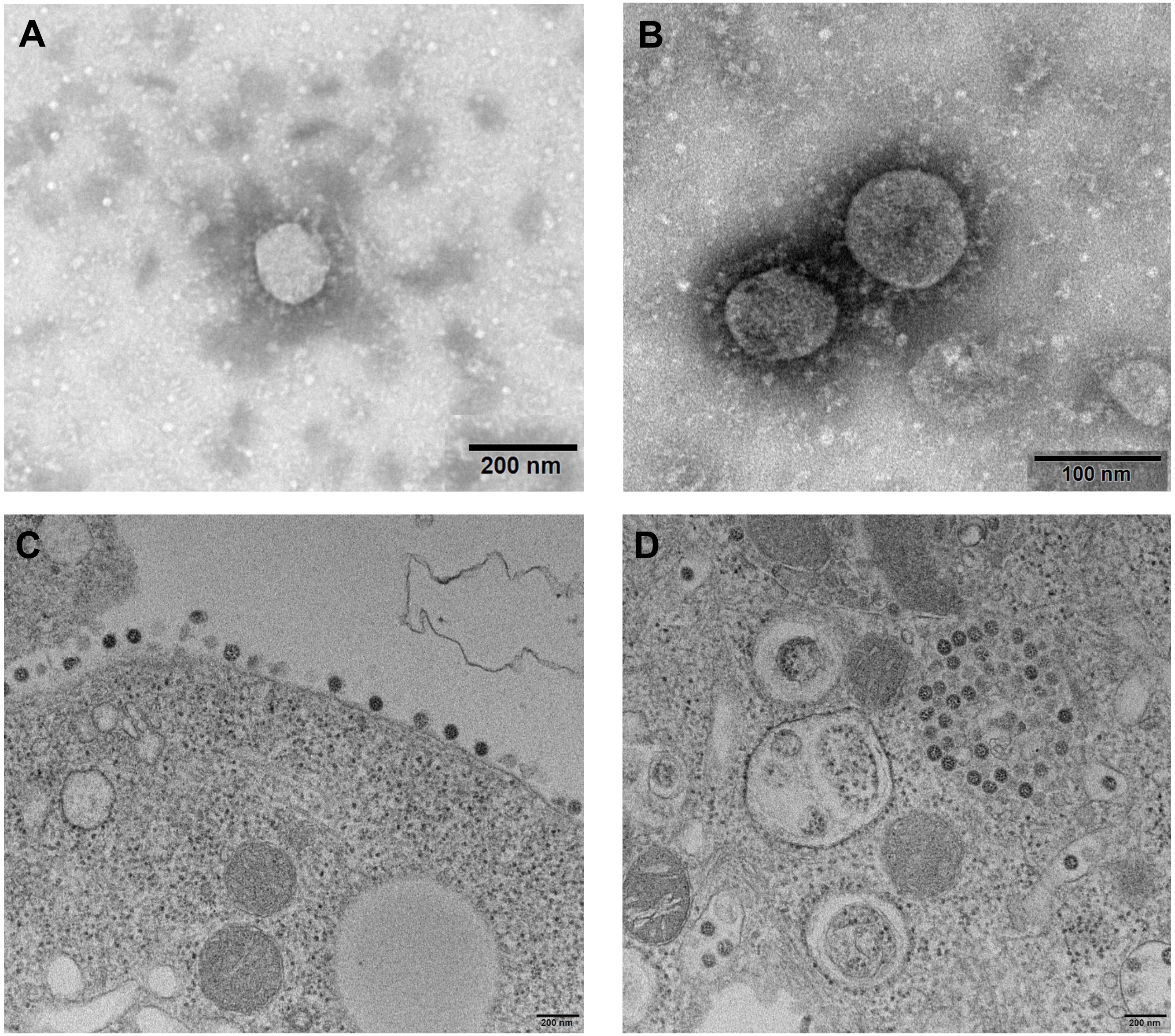
[d] The “original” means the sequencing results of the original environmental samples. The P1 and P3 mean the sequencing results of the cell supernatants from passage 1 and 3, respectively, during the virus isolation.

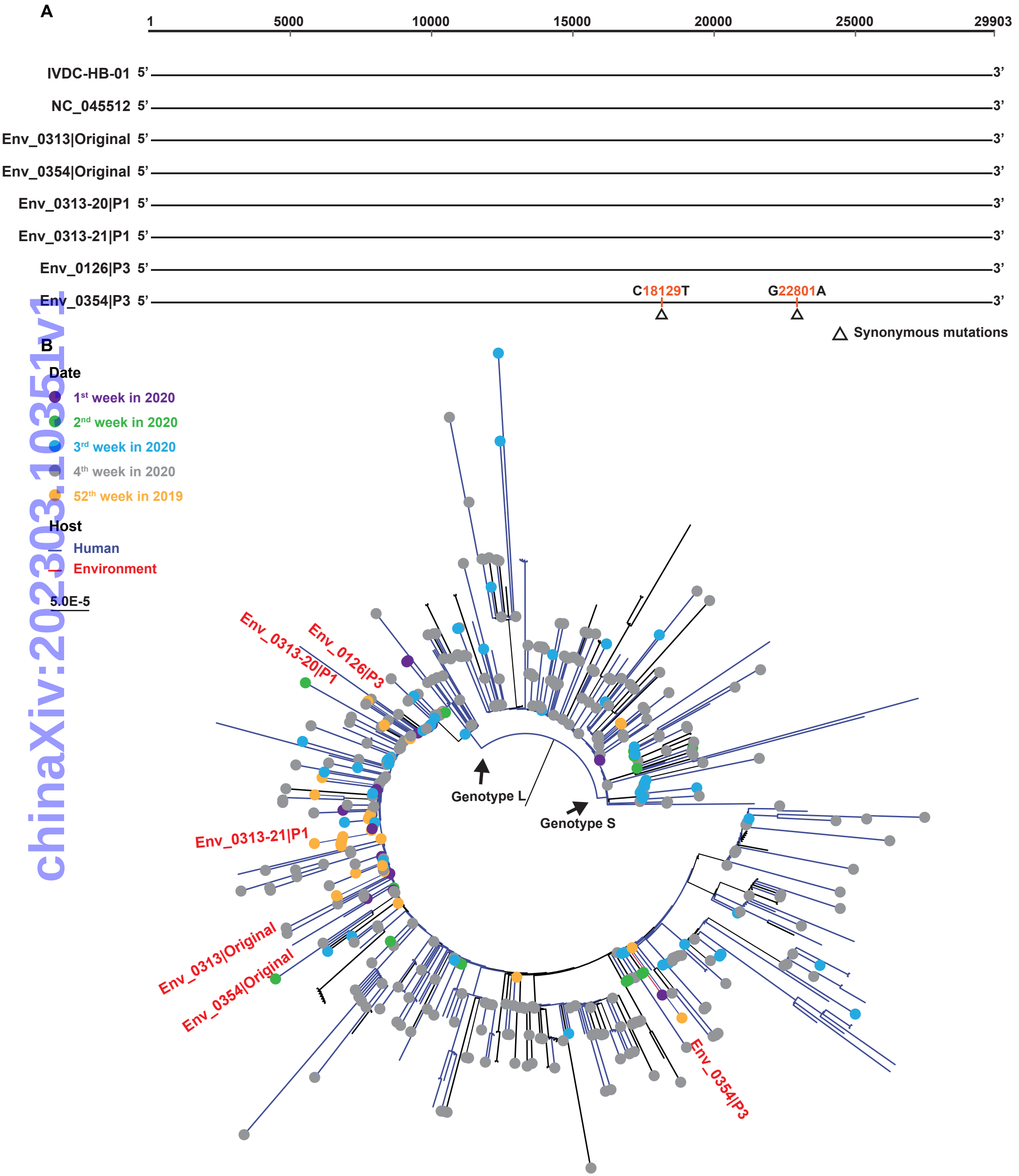
Extended Data Table 8. Summary of number of reads mapped to positions 8782 and 28144 in different samples[a].

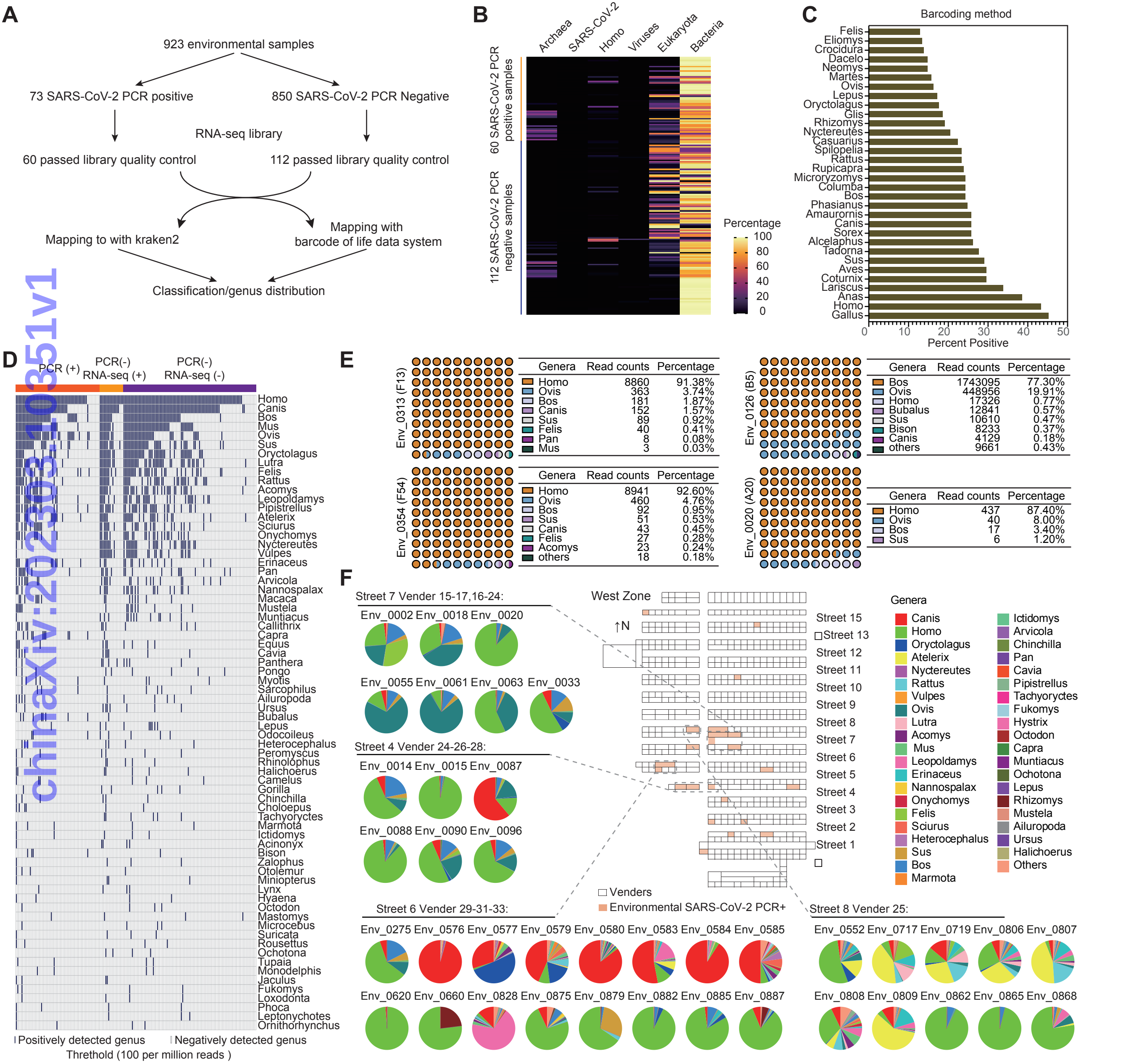
Sample ID	Lab Code	Position							
		8782				28144			
		A	T	C	G	A	T	C	G
Env_0101	A101	0	0	0	0	0	0	0	0
Env_0014	A14	0	0	0	0	0	0	0	0
Env_0015	A15	0	0	0	0	0	0	0	0
Env_0018	A18	0	0	0	0	0	0	0	0
Env_0020	A20	20	1044	24	0	3	19	1343	5
Env_0002	A2	0	0	0	0	0	0	0	0
Env_0033	A33	0	0	2	0	0	0	1	0
Env_0055	A55	0	0	0	0	0	0	0	0
Env_0061	A61	0	0	2	0	0	0	1	0
Env_0063	A63	0	0	0	0	0	0	0	0
Env_0087	A87	0	0	0	0	0	1	0	0
Env_0088	A88	0	0	0	0	0	2	0	0
Env_0090	A90	0	0	0	0	0	0	0	0
Env_0096	A96	0	0	0	0	0	0	0	0
Env_0138	B17	0	0	0	0	0	0	0	0
Env_0126	B5	0	0	86	0	0	85	0	0
Env_0213	D32	0	0	0	0	0	0	0	0
Env_0262	E48	0	0	0	0	0	0	0	0
Env_0275	E61	0	0	0	0	0	0	0	0
Env_0221	E7	0	0	0	0	0	0	0	0
Env_0400	F100	0	0	0	0	0	0	0	0
Env_0313	F13	6	13	403	4	11	644	6	6
Env_0333	F33	0	0	0	0	0	0	0	0
Env_0346	F46	0	0	1	0	0	0	0	0
Env_0354	F54	4	0	99	0	3	149	3	1
Env_0398	F98	0	0	20	0	0	3	0	0
Env_0509	G93	0	0	0	0	0	0	0	0

[a] These data are based on raw data in Extended Data Table 6.


A
↑N
West Zone
East Zone
Back Street
Street 15
Street 13
Street 12
Street 11
Street 10
Street 9
Street 8
Street 7
Street 6
Street 5
Street 4
Street 3
Street 2
Street 1
Stree
Xinhua Road
Fazhan Road
Venders
Live viruses were isolated
Environmental PCR +
Environmental PCR -
Sewers or sewerage wells PCR +
Sewers or sewerage wells PCR -
Domesticated wildlife products
Confirmed patient (N=1)
B
Environmental samples (NAT Positive)
Environmental samples (NAT Negative)
Animal samples (NAT Negative)
Number of samples
600
500
120
90
60
30
0
2020/1/1
2020/1/5
2020/1/10
2020/1/15
2020/1/20
2020/1/25
2020/1/30
2020/2/4
2020/2/9
2020/2/14
2020/2/19
2020/2/24
2020/2/29
2020/3/5
2020/3/10
2020/3/15
2020/3/20
2020/3/25
2020/3/30
Collection Date


A
200 nm
B
100 nm
C
200 nm
D
200 nm


A
1
5000
10000
15000
20000
25000
29903
IVDC-HB-01 5'
3'
NC_045512 5'
3'
Env_0313|Original 5'
3'
Env_0354|Original 5'
3'
Env_0313-20|P1 5'
3'
Env_0313-21|P1 5'
3'
Env_0126|P3 5'
3'
Env_0354|P3 5'
C18129T
G22801A
3'
Synonymous mutations
B
Date
1st week in 2020
2nd week in 2020
3rd week in 2020
4th week in 2020
52th week in 2019
Host
Human
Environment
5.0E-5
Env_0313-20|P1
Env_0126|P3
Env_0313-21|P1
Env_0313|Original
Env_0354|Original
Env_0354|P3
Genotype L
Genotype S



A
923 environmental samples
73 SARS-CoV-2 PCR positive
850 SARS-CoV-2 PCR Negative
RNA-seq library
60 passed library quality control
112 passed library quality control
Mapping to with kraken2
Mapping with barcode of life data system
Classification/genus distribution
B
Archaea
SARS-CoV-2
Homo
Viruses
Eukaryota
Bacteria
60 SARS-CoV-2 PCR positive samples
112 SARS-CoV-2 PCR negative samples
Percentage
100
80
60
40
20
0
C
Barcoding method
Felis
Eliomys
Crocidura
Dacelo
Neomys
Martes
Ovis
Lepus
Oryctolagus
Glis
Rhizomys
Nyctereutes
Casuarius
Spilopelia
Rattus
Rupicapra
Microryzomys
Columba
Bos
Phasianus
Amaurornis
Canis
Sorex
Alcelaphus
Tadorna
Sus
Aves
Coturnix
Lariscus
Anas
Homo
Gallus
0
10
20
30
40
50
Percent Positive
D
PCR (+)
PCR(-) RNA-seq (+)
PCR(-) RNA-seq (-)
Homo
Canis
Bos
Mus
Ovis
Sus
Oryctolagus
Lutra
Felis
Rattus
Acomys
Leopoldamys
Pipistrellus
Atelerix
Sciurus
Onychomys
Nyctereutes
Vulpes
Erinaceus
Pan
Arvicola
Nannospalax
Macaca
Mustela
Muntiacus
Callithrix
Capra
Equus
Cavia
Panthera
Pongo
Myotis
Sarcophilus
Ailuropoda
Ursus
Bubalus
Lepus
Odocoileus
Heterocephalus
Peromyscus
Rhinolophus
Halichoerus
Camelus
Gorilla
Chinchilla
Choloepus
Tachyoryctes
Marmota
Ictidomys
Acinonyx
Bison
Zalophus
Otolemur
Miniopterus
Lynx
Hyaena
Octodon
Mastomys
Microcebus
Suricata
Rousettus
Ochotona
Tupaia
Monodelphis
Jaculus
Fukomys
Loxodonta
Phoca
Leptonychotes
Ornithorhynchus
Positively detected genus
Negatively detected genus
Threthold (100 per million reads)
E
Env_0313 (F13)
Genera | Read counts | Percentage
Homo | 8860 | 91.38%
Ovis | 363 | 3.74%
Bos | 181 | 1.87%
Canis | 152 | 1.57%
Sus | 89 | 0.92%
Felis | 40 | 0.41%
Pan | 8 | 0.08%
Mus | 3 | 0.03%
Env_0126 (B5)
Genera | Read counts | Percentage
Bos | 1743095 | 77.30%
Ovis | 448956 | 19.91%
Homo | 17326 | 0.77%
Bubalus | 12841 | 0.57%
Sus | 10610 | 0.47%
Bison | 8233 | 0.37%
Canis | 4129 | 0.18%
others | 9661 | 0.43%
Env_0354 (F54)
Genera | Read counts | Percentage
Homo | 8941 | 92.60%
Ovis | 460 | 4.76%
Bos | 92 | 0.95%
Sus | 51 | 0.53%
Canis | 43 | 0.45%
Felis | 27 | 0.28%
Acomys | 23 | 0.24%
others | 18 | 0.18%
Env_0020 (A20)
Genera | Read counts | Percentage
Homo | 437 | 87.40%
Ovis | 40 | 8.00%
Bos | 17 | 3.40%
Sus | 6 | 1.20%
F
Street 7 Vender 15-17,16-24:
Env_0002 Env_0018 Env_0020
Env_0055 Env_0061 Env_0063 Env_0033
Street 4 Vender 24-26-28:
Env_0014 Env_0015 Env_0087
Env_0088 Env_0090 Env_0096
West Zone
↑N
Street 15
Street 13
Street 12
Street 11
Street 10
Street 9
Street 8
Street 7
Street 6
Street 5
Street 4
Street 3
Street 2
Street 1
Venders
Environmental SARS-CoV-2 PCR+
Street 6 Vender 29-31-33:
Env_0275 Env_0576 Env_0577 Env_0579 Env_0580 Env_0583 Env_0584 Env_0585
Env_0620 Env_0660 Env_0828 Env_0875 Env_0879 Env_0882 Env_0885 Env_0887
Street 8 Vender 25:
Env_0552 Env_0717 Env_0719 Env_0806 Env_0807
Env_0808 Env_0809 Env_0862 Env_0865 Env_0868
Genera
Canis
Homo
Oryctolagus
Atelerix
Nyctereutes
Rattus
Vulpes
Ovis
Lutra
Acomys
Mus
Leopoldamys
Erinaceus
Nannospalax
Onychomys
Felis
Sciurus
Heterocephalus
Sus
Bos
Marmota
Ictidomys
Arvicola
Chinchilla
Pan
Cavia
Pipistrellus
Tachyoryctes
Fukomys
Hystrix
Octodon
Capra
Muntiacus
Ochotona
Lepus
Rhizomys
Mustela
Ailuropoda
Ursus
Halichoerus
Others